

MBL/WHOI
0 0301 0010818 9



FEB 17 1959

ELECTRONIC APPARATUS FOR
BIOLOGICAL RESEARCH

ELECTRONIC APPARATUS

FOR

BIOLOGICAL RESEARCH

P. E. K. DONALDSON, M.A.

with contributions by

J. W. L. Beament, M.A., Ph.D.

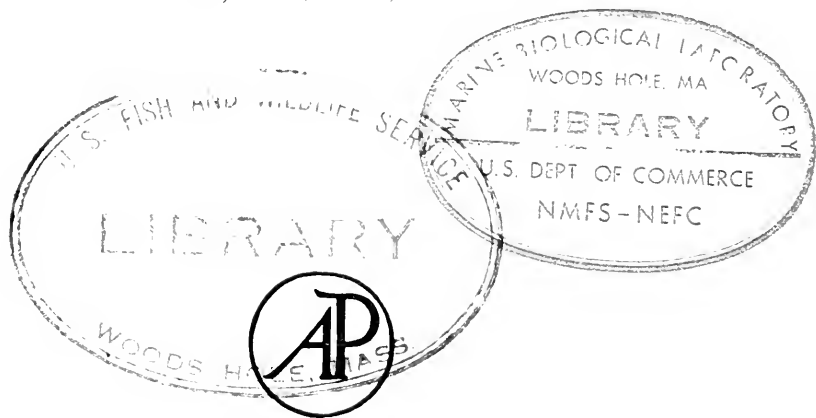
F. W. Campbell, M.A.

D. W. Kennard, M.B., B.S., M.R.C.S., L.R.C.P., Ph.D.

R. D. Keynes, M.A., Ph.D.

K. E. Machin, M.A., Ph.D., A.M.I.E.E.

I. A. Silver, M.A., Ph.D., M.R.C.V.S.



NEW YORK

ACADEMIC PRESS INC., PUBLISHERS

LONDON

BUTTERWORTHS SCIENTIFIC PUBLICATIONS

1958

BUTTERWORTHS PUBLICATIONS LTD
88 KINGSWAY, LONDON, W.C.2

U.S.A. Edition published by
ACADEMIC PRESS INC., PUBLISHERS
111 FIFTH AVENUE
NEW YORK 3, NEW YORK

5-126404

©

The Author and Contributors named on the title page
1958

Printed in Northern Ireland at The Universities Press, Belfast

CONTENTS

	<i>Page</i>
Foreword	<i>ix</i>
Preface	<i>xi</i>
PART I : THEORY	
1 INTRODUCTION	3
2 RESISTANCES	6
3 RESISTANCES AND CAPACITANCES	25
4 INDUCTANCES AND RESISTANCES	54
5 INDUCTANCES, CAPACITANCES AND RESISTANCES	72
6 DIODE CIRCUITS	96
7 SOFT VALVES	118
8 HARD VALVES	132
9 INTERSTAGE COUPLING	149
10 DIFFICULTIES WITH SINGLE-SIDED AMPLIFIERS	159
11 NEGATIVE VOLTAGE FEEDBACK AND THE STABILIZED GAIN AMPLIFIER	164
12 DOUBLE-SIDED AMPLIFIERS	176
13 TUNED AMPLIFIERS	196
14 SINE WAVE OSCILLATORS	208
15 SQUARE WAVE OSCILLATORS	226
16 TRIGGERED PULSE GENERATORS	235
17 NOISE	253
18 THE CATHODALLY SCREENED CATHODE FOLLOWER	259
GRAPHS	263

CONTENTS

PART II: PRACTICE

19	BATTERIES	283
20	RESISTORS	291
21	CAPACITORS	297
22	CHOKES AND TRANSFORMERS	305
23	NON-THERMIONIC DIODES	315
24	VALVES	319
25	SUNDRIES	321
26	FORMS OF MECHANICAL CONSTRUCTION	327
27	TOOLS AND WORKSHOP FACILITIES	329

PART III: TRANSDUCERS, ELECTRODES AND INDICATORS

28	LIGHT SOURCES AND DETECTORS <i>F. W. Campbell</i>	333
29	MEASUREMENT AND CONTROL OF TEMPERATURE <i>J. W. L. Beament</i>	383
30	MEASUREMENT AND CONTROL OF HUMIDITY <i>J. W. L. Beament</i>	413
31	ASSAY OF RADIOACTIVITY <i>R. D. Keynes</i>	419
32	VISUAL INDICATORS <i>P. E. K. Donaldson</i>	446
33	TRANSDUCERS <i>K. E. Machin</i>	471
34	RELAYS AND RELATED MECHANISMS <i>K. E. Machin</i>	510
35	GLASS MICROCAPILLARY ELECTRODES USED FOR MEASURING POTENTIAL IN LIVING TISSUES <i>D. W. Kennard</i>	534
36	OTHER ELECTRODES <i>I. A. Silver</i>	568

CONTENTS

PART IV: COMPLETE APPARATUS

37	POWER PACKS	585
38	STIMULATORS	602
39	BIOLOGICAL AMPLIFIERS	616
40	METHODS OF RECORDING	638
41	TIMERS, COUNTERS AND RATE MEASUREMENT	643
42	LAYOUT AND THE CONTROL OF INTERFERENCE	656
43	FAULTS	665
44	DESIGN PROCEDURE	671
45	TRANSISTORS	677
	Index	711

FOREWORD

This book will, I hope, prove a valuable technical handbook for those engaged in the use of electronic techniques in biology. I think its significance and value should be much more than as a guide to technique. It is easy for users of elaborate techniques to come to regard them as tools they need not bother to understand, and the reliability of modern apparatus is such as to encourage this attitude; it is, however, a disastrous one for a scientist who interprets the results obtained and who must be the master and not the slave of his technology. This book enables him to remain master despite complex techniques.

For over 25 years research students who have come to Cambridge to learn electrophysiology have been made to construct with their own hands elementary biological electronic apparatus before being allowed to work with ready-made recording systems. Even those who at the time slightly resented being returned to the electronic kindergarten, thought it in retrospect invaluable to have mastered the details, though experts were available to manage the apparatus thereafter. The present book carries on this tradition and enables a biologist to understand the basis, and thus the capabilities and limitations, of the methods available to him.

Donaldson, working and writing in a laboratory which has been a main centre of electrophysiology for 50 years and in the forefront of the biological use of electronic devices for more than 30, has presented not only a compendium of the formal bones of these techniques, but also incorporated many of those little points of know-how that come from long practical experience.

I think this book will be valuable both to the beginner entering this field and also to the more experienced, for its comprehensiveness makes it a reference work of electronics in biology.

BRYAN MATTHEWS

*Physiological Laboratory,
Cambridge*



PREFACE

In offering this book under the present title I am aware that of many possible criticisms, one might run as follows: 'Biological research is a tremendous subject, continuously advancing on many fronts. How can you, with five years' experience solely in a physiological laboratory, write with authority about instrumentation for a field as broad as your title suggests?'

The book might have been planned in two ways: first, the extent of biological research could have been surveyed and classified. Under each heading, applications for electronic apparatus could have been discussed, then particular examples of suitable instruments taken and constructional details given. Put together along these lines, the book would almost certainly have been both long and repetitious. In addition, a newcomer to electronics who had successfully built a piece of equipment from the instructions given might neither understand how it worked nor be able to modify it to extend its capabilities, or even know what to do when it went wrong.

The repetition just alluded to gives the clue to an alternative structure; that which has been adopted in fact. Electronic techniques, though diverse in application, are actually rather few in number; the main difference between apparatus used by workers in widely separated research fields lies often less in the electronic gear proper than in the transducing devices employed; that is, on the devices that link the electronic circuits with the remainder of the experimental set-up. Thus Worker A, interested in the movement of radioactive isotopes in biological tissue, and Worker B, who wishes to count the passage of action potentials along a nerve, may each employ rather similar valve amplifiers and electronic counters; but A's transducer is a Geiger-Müller tube, whereas B's is an electrode system of silver wires.

In this book, then, the spotting of applications is left largely to the reader; the emphasis is upon techniques. Part I is concerned with theory; beginning with the behaviour of electrical circuit elements, first singly and then in groups, networks of increasing complexity are considered, up to what might be called 'functional units', e.g., a stage of amplification, a filter, or an oscillator. Part II is practical, and is mostly concerned with electronic components, their ratings, limitations and correct use. I have here to acknowledge the kindness of the following manufacturers, who freely supplied samples of their products to make possible the plates in this Part: Painton and Co., Ltd., for Multicon connectors and high-stability resistors; Welwyn Electrical

PREFACE

Laboratories for wirewound resistors; Westinghouse Brake and Signal Co., for copper-oxide rectifiers and diodes; A. H. Hunt (Capacitors) Ltd., for metallized paper capacitors; Ferranti Ltd., and the British Thomson-Houston Co. Ltd., for junction diodes; Standard Telephones and Cables Ltd., for a selenium diode; and Mullard Ltd., for 'Ferroxcube' cores and bobbins, and for permission to reproduce parts of their 'Ferroxcube' handbook.

Part III is devoted to transducers in the broad sense; that is, it includes chapters on electrodes and on indicators. This part has been written largely by colleagues whom I regard as experts in their several fields, and for whose willing co-operation I must record my appreciation and thanks.

In Part IV, I have tried to draw together all that has gone before, to discuss complete pieces of apparatus, indicating how they are assembled from 'functional units', and how to design, use, and maintain them.

With the needs of the biologist new to electronics in mind, my aim has been to provide in one volume both a textbook which begins with first principles and a work of reference. Naturally, no book alone can replace a proper course of technical instruction with appropriate practical work, but if the reader will 'lash-up' for himself, and experiment with, some of the circuits given as he reads about them; and if he can find time to follow the technical magazines, paying as much attention to the advertisements as to the editorial matter (one learns much about components in this way), then it is my hope that he will be able to design, and construct for himself, electronic apparatus which will serve him well.

Finally, I wish to thank Professor Sir Bryan Matthews for his support and encouragement; my colleagues in this Laboratory for many valuable discussions; Drs. R. H. Adrian and G. S. Brindley, for criticizing and checking the text; Miss Alison Howlett, who took the photographs; and my Wife, who performed many hundreds of calculations involved in plotting out the graphs.

P. E. K. D.

*Physiological Laboratory,
Cambridge, 1958*

PART I
THEORY



INTRODUCTION

Pieces of electronic apparatus are complexes of 'active' and 'passive' elements. Active elements are to be regarded as sources of electrical energy, whilst the passive elements either consume energy or modify it in some way. For an electrical circuit to be possible at least one element of each kind must be present. In developing a systematic theory of such circuits a problem of the chicken-and-the-egg variety arises—how to study one set of elements without having first discussed the other.

The important active elements, in practical form, are of two kinds: (1) those which actually produce electrical energy—batteries, the supply mains and certain transducers; (2) those which only control the flow of energy derived from elsewhere, but which are treated as if they were themselves the source of it—valves, transistors, and the remainder of the transducers.

The passive elements are resistance, capacitance, and inductance, properties realized either in a specific component, e.g. a *resistor*, which is a device for exhibiting the property of *resistance*, or otherwise, as in the resistance of an electrophysiological preparation as measured between two electrodes. It may be remarked in passing that such a preparation is to be regarded as an active element of the first kind at the site of recording, but as passive at a site of stimulation.

Since it is not possible to get very far discussing the active devices without introducing some passive ones it is usual to begin by considering the latter, and this we shall do, beginning in the next chapter. To investigate the properties of the passive elements, however, both singly and in combination, we need some ideal, generalized kinds of active element which we shall simply call 'generators'. A 'load' is to be understood as any passive element or arrangement of such elements into which these generators may work.

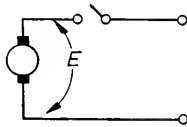


Figure 1.1

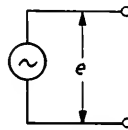


Figure 1.2

The ideal generators are:

(1) The direct voltage, constant-voltage generator (*Figure 1.1*). This is supposed to be a source of constant e.m.f. E , which will maintain this e.m.f. however large or small the current it supplies.

(2) The alternating voltage, constant-voltage generator (*Figure 1.2*). A source of e.m.f. $e = E \sin \omega t$, which, like *Figure 1.1*, continues to generate this e.m.f. whatever the load condition. The title reads rather paradoxically until it is appreciated that the constancy referred to is with respect to load

INTRODUCTION

rather than to time. For brevity we shall also call *Figure 1.2* the ‘constant-voltage alternator’.

(3) The direct current, constant-current generator (*Figure 1.3*). This is supposed to force a constant direct current I through the load, whatever the nature of the load.

(4) The alternating current, constant-current generator (*Figure 1.4*). This forces a current $i = I \sin \omega t$ through the load, whatever the nature of the load. We shall also call this the ‘constant-current alternator’.

Notice that the direct voltage and current generators are drawn with a switch. The former is ‘on’ when the switch is closed, and the latter is ‘on’ when the switch is open. The reason for the presence of these switches is important, and necessitates a rather long digression.

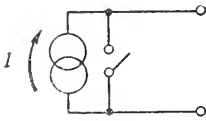


Figure 1.3

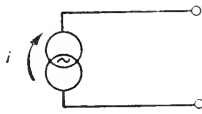


Figure 1.4

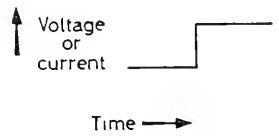


Figure 1.5

Electrotechnology may be divided into two kinds, ‘Power’ engineering and ‘Control’ engineering. In the former the commodity is electrical energy, and the object is to convey this in a continuous and satisfactory manner from the source to the consumer: the abiding consideration is efficiency. In control engineering the commodity is *change*, information, a message or *signal*. Power efficiency is usually unsought here; the important thing is to minimize any distortions which the signal may undergo in its passage through the apparatus.

Electronics in biological research is a department of control engineering. The techniques allow information about biological happenings to be transferred to the observer, and there is a sense in which they allow signals to be passed from the observer to the biological system, e.g. as in delivering a controlled stimulation. It follows that in descriptions of the performance of such apparatus, the amount of distortion which it introduces recurs frequently.

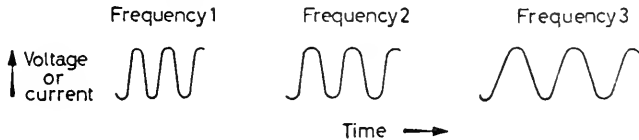
Distortion may be placed upon a quantitative basis by considering the difference between the ideal response and the actual response obtained, but this cannot be stated for every possible kind of signal, since signals are usually innumerable. The best that can be done is to consider the distortions arising when apparatus receives special signals of a simple and artificial sort.

In electrobiology, signals are mainly of the continuously variable kind, that is, the apparatus is analogic rather than digital; a voltage or current represents the extension of a muscle, the light reflected by a retina, or a membrane potential. Our test signals, which will allow us to estimate the performance of our apparatus, will therefore be continuous, single-valued functions of time. In practice, the two commonest are: (1) the step-function (*Figure 1.5*) in which the input to the apparatus changes abruptly from one

REFERENCES

steady value to another; (2) the sine wave of constant amplitude and variable frequency (*Figure 1.6*).

The apparatus performance may then be described in terms of the *transient response* to the application of the step-function, with the introduction of terms such as 'overshoot', 'lag', 'time constant', or in terms of the *frequency response* or *steady-state response* to the sine wave, describing the amplitude and relative phase of the output as the frequency of the input—of constant amplitude—is varied. Phrases such as 'x decibels down at y cycles per second' are part of steady-state jargon.



Transient response analysis and steady-state analysis, then, are the methods by which the behaviour of component networks in control engineering are discussed. Historically the two analyses spring from rather different engineering fields; the former from radar and television, the latter from telephone engineering and radio. Radar and television are 'pulse' techniques, interested in angular waveforms like the step-function, whereas sound waveforms are smoother affairs, more obviously composed of a number of sine waves. In electronics as applied to biological research both kinds of waveform occur and both analyses are relevant. Both analyses are really describing the same thing, being connected by the Fourier transform:

$$f(t) = \lim_{A \rightarrow \infty} \int_{-A}^{+A} e^{j\omega t} F(\omega) d\omega$$

The step-function response $f(t)$ is composed of the sum of a large number of component sine waves, but the numerical production of the transient response of a network whose steady-state response is known, or vice versa, is apt to be a tedious business^{1,2,3}.

In this book the emphasis is upon steady-state response, though transient response is discussed in the simpler and more ubiquitous cases. The evaluation of steady-state response is easier mathematically, elementary algebra being all that is required, whereas the solution of the most simple transient response problem involves a differential equation. Furthermore, in the writer's opinion, steady-state response of actual apparatus is easier to measure.

The reason for the introduction of the four kinds of ideal generator is now evident; the alternating voltage and current generators are for investigating steady-state behaviour; the direct generators are for obtaining a step-function of voltage or current.

REFERENCES

- ¹ RODDAM, T. Calculating transient response *Wireless World* 63 (1952) 292
- ² LUDBROOK L. C. Step to frequency response transforms *Electron. Engng.* 26, (1954) 27
- ³ JAWORSKI, Z. E. Empirical transient formulae *Electron. Engng.* 26 (1954) 397

RESISTANCES

ELEMENTARY ELECTRICAL CIRCUIT

If a constant-voltage generator of e.m.f. E is connected to a resistance R , the current which flows (*Figure 2.1*) is

$$I = \frac{E}{R}$$

I is in amps if E is measured in volts and R in ohms. This is the relationship usually called Ohm's law, though in fact Ohm only said that I would be proportional to E if R were kept constant.

If a generator of constant current I is connected to a resistance R (*Figure 2.2*) a difference of potential V is caused to appear across the resistance. $V = IR$, therefore $I = V/R$. V is in volts if I is measured in amps and R in ohms.

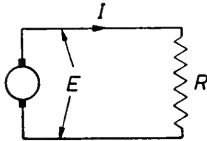


Figure 2.1

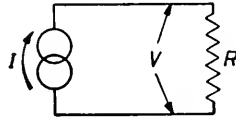


Figure 2.2

Though E and V are both voltages the distinction between them is theoretically important, since E is a cause and V an effect. Whilst it is possible to preserve the distinction when discussing very simple circuits it is not so easy in more complicated cases; this is because electronic circuits are arranged in causal chains. Thus a voltage which is V for the n th link in the chain will be E for the $n + 1$ th. In such cases we shall simply speak of 'the voltage' between such-and-such a pair of points, using the symbol V .

In either *Figure 2.1* or *Figure 2.2* the quantity of electricity passing any point in the circuit in time t is

$$Q = It \quad (\text{coulombs, amps, seconds})$$

The rate of expenditure of energy in the resistance is called the 'power' and is

$$P = VI \text{ or } I^2R \text{ or } \frac{V^2}{R} \quad (\text{watts, volts, amps, ohms})$$

The expenditure of energy in time t is

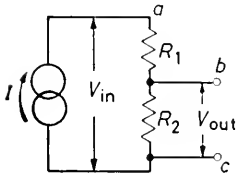
$$W = Pt \quad (\text{joules, watts, seconds})$$

The potential-divider and potentiometer

Resistances connected as in *Figure 2.3a* are said to be in 'series', and the total resistance in the circuit is $R_1 + R_2$. In electronics we often wish to reduce a potential difference to some fraction, either fixed or variable. If a current I flows through R_1 and R_2 , it produces a potential difference of IR_1 across R_1 and IR_2 across R_2 , making a total of $I(R_1 + R_2)$ across the pair. In the *potential divider* the total potential difference $I(R_1 + R_2)$ is thus divided into two fractions of which one, IR_2 , is made use of at the output terminals bc . The *transmission factor* of the device is

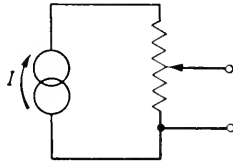
$$\frac{V_{out}}{V_{in}} = \frac{IR_2}{I(R_1 + R_2)} = \frac{R_2}{(R_1 + R_2)}$$

If the transmission factor is to be made variable, the device is called a *potentiometer* (*Figure 2.3b*) in which a 'tap' connection can be moved up and down a single resistance, dividing it into two parts. $(R_1 + R_2)$ is now a constant and the transmission factor is proportional to R_2 .



(a)

Figure 2.3



(b)

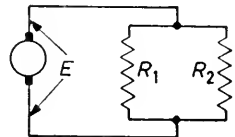


Figure 2.4

Resistances in parallel

Resistances connected as in *Figure 2.4* are said to be connected in parallel or shunt. The current in R_1 is E/R_1 and in R_2 is E/R_2 . Therefore, with resistances in parallel the current divides inversely as the ratio of the resistances. The total current I_{tot} is $E/R_1 + E/R_2$. The effective resistance of the combination is thus

$$R_{eff} = \frac{E}{I_{tot}} = \frac{E}{\frac{E}{R_1} + \frac{E}{R_2}}$$

Dividing top and bottom by E , we find that for two resistances in parallel

$$R_{eff} = \frac{I}{\frac{I}{R_1} + \frac{I}{R_2}}$$

and in general for any number of resistances

$$R_{eff} = \frac{I}{\frac{I}{R_1} + \frac{I}{R_2} + \dots + \frac{I}{R_n}}$$

RESISTANCES

A more convenient form in the case of two resistances is

$$R_{\text{eff}} = \frac{R_1 R_2}{R_1 + R_2}$$

REAL GENERATORS

We are now in a position to consider real generators. The definition of a constant-voltage generator implies that the 'internal resistance' is zero, and that of a constant-current generator that the internal resistance is infinite. Real generators have internal resistances between zero and infinity, and may be represented either: by an ideal constant-voltage generator in series with the internal resistance r (Figure 2.5a), or by an ideal constant-current generator in parallel with the internal resistance r (Figure 2.5b).

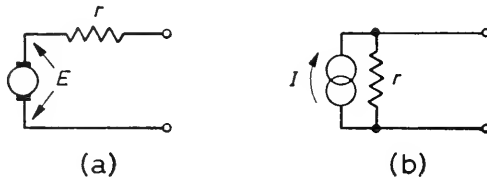


Figure 2.5

If I in Figure 2.5b is equal to E in Figure 2.5a the two diagrams represent the same thing, for if the P.D. across the terminals of the latter is measured with no load resistance connected, the open-circuit voltage is clearly E . If the terminals are short circuited, the short-circuit current is E/r . If a load resistance R be connected, the current through it is $E/(r + R)$ and the voltage across it will be $E\{R/(r + R)\}$. Similarly, in Figure 2.5b the open-circuit voltage is I going through r , which is I . The short-circuit current is just I , since it will all go through the short-circuit rather than through r . The voltage across a load resistance R will be I times the effective resistance of R in parallel with r , which is $I(rR)/(r + R)$, and the current through the load will be this voltage divided by R , i.e.

$$\frac{I \frac{rR}{r + R}}{R} = I \frac{r}{r + R}$$

Summarizing in tabular form:

	Figure 2.5a	Figure 2.5b
Open-circuit voltage	E	I
Short-circuit current	E/r	I
Voltage across load R	$E\{R/(r + R)\}$	$I\{(rR)/(r + R)\}$
Current through load R	$E/(r + R)$	$I/r/(r + R)$

REAL GENERATORS

Remembering that the relationship which equates the two arrangements is $E = Ir$, if Ir be substituted into the first column wherever E appears, the two columns will be found to be identical. The two representations are therefore equivalent.

Having established that a real generator may be represented in either of two ways, it remains to decide which way is better in a particular case. Consider a real generator of internal resistance r , and suppose it has a measured open-circuit voltage E . Let us plot the output voltage and current as a load resistance is made less than, equal to, and greater than r . The relevant equations are, as we have seen, $V_{\text{out}} = E\{R/(r + R)\}$ and $I_{\text{out}} = E/(r + R)$ (Graph 1).

Clearly when $R \gg r$, the output voltage is rather constant but the current is not. And when $R \ll r$, the output current is rather constant but the voltage is not. Therefore,

If a real generator is feeding a load of resistance much greater than its own, the ideal constant-voltage generator symbol is appropriate, with the internal resistance in series. If a real generator is feeding a load of resistance much lower than its own, the ideal constant-current generator symbol is appropriate, with the internal resistance in shunt.

If the load resistance and generator resistance are of the same order, either representation will do, but the constant-voltage symbol seems to be commoner. A real generator having $r \ll R$ will be said to be of the constant-voltage type. A real generator having $r \gg R$ will be said to be of the constant-current type.

Maximum power transfer theorem

In Figure 2.6 we have a real generator feeding a load. An important problem is to decide what value of R extracts maximum power from the generator terminals.

The load current is $E/(r + R)$, so the load power, P , is $I^2R = \{E^2R\}/\{(r + R)^2\}$. When the load power is a maximum, the rate of change of load power with load resistance will be zero, i.e.

$$\frac{dP}{dR} = 0$$

$$\text{but } \frac{dP}{dR} = E^2 \left\{ \frac{-2R}{(r + R)^3} + \frac{1}{(r + R)^2} \right\}$$

$$\text{whence } R = r.$$

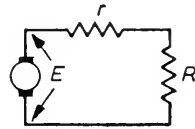


Figure 2.6

A load is said to be *matched* to a generator, and takes maximum power from it, when the load resistance is equal to the generator internal resistance. This is the maximum power transfer theorem. It does *not* mean that a generator delivers maximum power to a given load when the generator internal resistance is made equal to the load.

Easily solved resistance networks

The rules for resistances in parallel and in series can be used to solve completely a certain class of resistance network, which may be quite elaborate.

RESISTANCES

For example, the effective resistance of the network in *Figure 2.7* is easily found, as follows.

We first combine the resistances from *A* to *B* into one of $3 + 1.5$ ohms = 4.5 ohms. Similarly, between *C* and *D* we have $6 + 3 = 9$ ohms. The network simplifies to *Figure 2.8*. We now consider the 4.5 ohms and 9 ohms

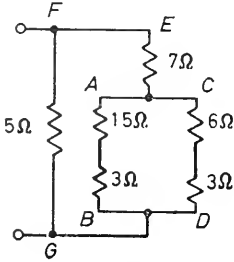


Figure 2.7

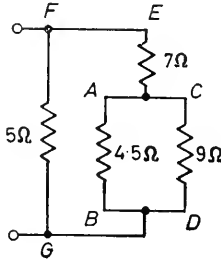


Figure 2.8

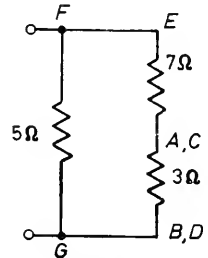


Figure 2.9

resistances in parallel; they are equivalent to a single resistance of $(9 \times 4.5)/(9 + 4.5) = 3$ ohms (*Figure 2.9*). Next we combine the newly found 3 ohms with the 7 ohms yielding *Figure 2.10*. Finally we are left merely with 5 ohms in parallel with 10 ohms, which reduces to a single equivalent resistance of $(5 \times 10)/(5 + 10) = 3.3\bar{3}$ ohms (*Figure 2.11*).

To solve all the voltages and currents which flow in *Figure 2.7* when a generator is connected across the terminals, one quick way is to assume part

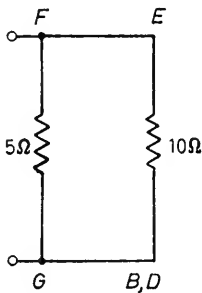


Figure 2.10

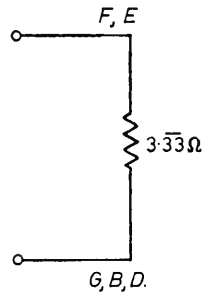


Figure 2.11

of the answer; e.g. that the current in *AB* is one amp downwards. Since the currents divide inversely as the ratio of parallel resistances, the current *CD* must be $\frac{1}{2}$ amp downwards. Thus the current in *E-AC* is $1\frac{1}{2}$ amps downwards. In *Figure 2.10*, if we have $1\frac{1}{2}$ amps flowing from *E* to *BD*, the current down *F-G* must be 3 amps, and the total current supplied to the network must be $4\frac{1}{2}$ amps. If we are told that the generator *actually* supplies 13 amps to the network, then the actual currents flowing in the various parts of the network are found merely by multiplying the assumed values by $13/4\frac{1}{2}$. Finally the potential differences across the resistances are worked out by multiplying the current through each resistance by its value in ohms.

REAL GENERATORS

The whole process is much quicker to do than to describe. If, instead of being told the total current supplied to the network we had been told the e.m.f. to which it had been connected, we would have worked out the total current by dividing the effective resistance of the network, 3.33 ohms, into the e.m.f.

The class of network which can be solved along these lines may be defined in the following way. If for every point where a current divides into two fractions there is another point at which the same two fractions are reunited, then the network may be solved by the rules for series and parallel resistances. Thus, *Figure 2.7* is soluble because the currents which divide at *F* meet again at *G*, and those which part company at *AC* coalesce once more at *BD*.

A network which does not fulfill these conditions is the unbalanced Wheatstone bridge (*Figure 2.12*). Current divides at *A*, flows down the two arms of the bridge and meets at *B*. If current is flowing along *C-D*, clearly the two currents leaving *A* are not in the same ratio as those meeting at *B*. The classical way of solving this problem, and its elaborations, is to resort to Kirchhoff's laws, setting up and solving a number of simultaneous equations. We shall not do this, but have recourse instead to two powerful weapons which follow from Kirchhoff's laws. They are the theorem of Thévenin and the star-delta transformation.

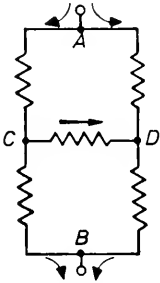


Figure 2.12

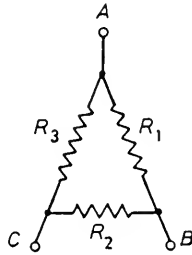


Figure 2.13

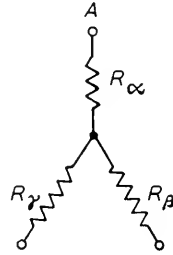


Figure 2.14

The star-delta transformation

If three resistances have the configuration of *Figure 2.13* they are said to form a 'mesh' or 'delta'. It can be shown that an equivalent arrangement is that of *Figure 2.14*, called a 'star', provided the resistances comprising the star are related to the mesh resistances in the following manner:

$$R_{\alpha} = \frac{R_1 R_3}{R_1 + R_2 + R_3}$$

$$R_{\beta} = \frac{R_2 R_1}{R_1 + R_2 + R_3}$$

$$R_{\gamma} = \frac{R_3 R_2}{R_1 + R_2 + R_3}$$

RESISTANCES

To perform the reverse process, that of converting a star into the equivalent delta, the equations are:

$$R_1 = R_x + R_\beta + \frac{R_x R_\beta}{R_y}$$

$$R_2 = R_\beta + R_\gamma + \frac{R_\beta R_\gamma}{R_x}$$

$$R_3 = R_\gamma + R_x + \frac{R_\gamma R_x}{R_\beta}$$

Example of application—Figure 2.15 shows an unbalanced Wheatstone bridge. It is required to find the effective resistance between terminals *A* and *B*.

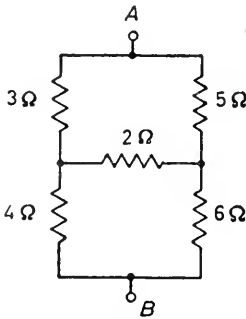


Figure 2.15

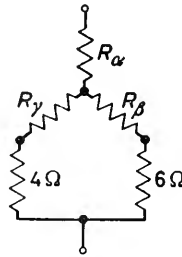


Figure 2.16

Notice that a delta is formed by the 3, 2 and 5 ohms resistances, and another by the 4, 2 and 6 ohms resistances. Choose one, say the first, and convert to the equivalent star (Figure 2.16).

$$R_\alpha = \frac{5 \times 3}{5 + 2 + 3} = 1.5 \Omega$$

$$R_\beta = \frac{5 \times 2}{5 + 2 + 3} = 1 \Omega$$

$$R_\gamma = \frac{2 \times 3}{5 + 2 + 3} = 0.6 \Omega$$

The remainder of the problem is quickly solved by the rules for series and parallel resistances.

Thévenin's theorem

This theorem is for finding the current in any member of a network of resistances being supplied from a generator. It may be stated as follows: If the resistance in question, *R*, be imagined to be removed, the remainder of the network may be represented by a constant voltage generator of e.m.f. *E'* in series with a resistance *r'*, where *E'* is the open-circuit voltage appearing across the terminals initially joined by *R*, and *r'* is the effective resistance 'looking into' these terminals. Thus, an equivalent circuit for the network with *R* in place is just (Figure 2.17) and the current through *R* is $E'/(R + r')$.

REAL GENERATORS

Example of application—Find the current in the 2 ohms resistance of the bridge circuit in *Figure 2.15* if a constant-voltage generator of e.m.f. 10 volts is connected across *A-B*.

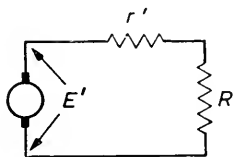


Figure 2.17

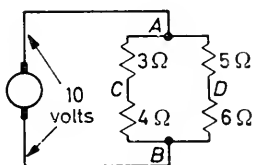


Figure 2.18

First, remove the 2 ohms resistance (*Figure 2.18*). We now have two potential dividers and the bridge can be solved by the rules for series resistances:

The resistance in the path *A-C-B* is 7 ohms;
 therefore the current in the path *A-C-B* is $10/7$ amps;
 therefore the potential difference *A-C* is $3 \times (10/7)$ volts.

By a similar process the P.D. *A-D* is $5 \times (10/11)$ volts. So the P.D. *C-D* is $\{5 \times (10/11)\} - \{3 \times (10/7)\}$ and this is our E' . To find r' , bearing in mind

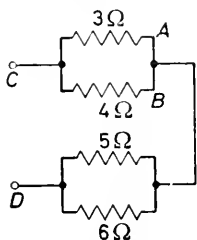


Figure 2.19

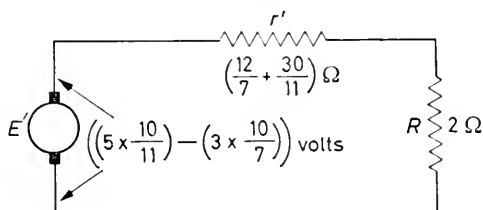


Figure 2.20

that the generator has no internal resistance, we re-draw the bridge as in *Figure 2.19*. 'Looking in' at *C-D*, we see a resistance of 3 ohms in parallel with 4 ohms, plus 5 ohms in parallel with 6 ohms.

$$r' = \frac{3 \times 4}{3 + 4} + \frac{5 \times 6}{5 + 6} = \frac{12}{7} + \frac{30}{11}$$

The equivalent circuit is therefore *Figure 2.20* and the current through R may be quickly found.

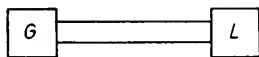


Figure 2.21



Figure 2.22



Figure 2.23

Potentiometer control of power

In *Figure 2.21* we have a generator G feeding some kind of load L . We often wish to control the amount of power being supplied to L . If G is of the constant-voltage type, such as the electric supply mains, and L is, say, a lamp, an appropriate method would be a variable series resistance (*Figure 2.22*) or 'rheostat' to 'drop' some of the voltage. Similarly, if G is of the

RESISTANCES

constant-current type, control may be had by a parallel variable resistance to 'bleed' some of the current (*Figure 2.23*). (Notice that no control is achieved by attempting to bleed a constant-voltage generator, or by connecting a variable dropping resistance in series with a constant-current generator.) If the generator is approximately matched to the load either method may be used, but the variable series resistance is to be preferred since it reduces the power to the load by reducing the power supplied by the generator, whereas a variable bleeder works the other way round, and is therefore rather uneconomical.

A difficulty with the variable series resistance is that, to reduce progressively the power supplied to the load to zero, the resistance must increase smoothly to infinity, which is usually rather difficult to arrange. The problem can be overcome by arranging the resistance as a potentiometer (*Figure 2.24*) when obviously the load power is zero when the slider is at the bottom. When the slider is at the top some power is delivered to the load and some is dissipated in the potentiometer. The latter fraction is small if the potentiometer resistance is high.

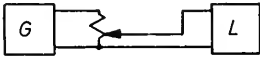


Figure 2.24

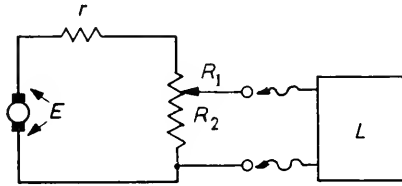


Figure 2.25

We saw earlier that if no load is connected to a potentiometer the output voltage is proportional to R_2 . If the potentiometer is of the 'linear' variety, R_2 is proportional to shaft rotation, so

$$\text{output voltage} \propto \text{shaft rotation}$$

When a load is connected this is no longer true. In *Figure 2.25*, if we disconnect the load and apply Thévenin's theorem to the output terminals of the potentiometer we see an open-circuit e.m.f., E' , of

$$E \frac{R_2}{R_1 + R_2 + r}$$

in series with a resistance

$$r' = \frac{R_2(R_1 + r)}{R_1 + R_2 + r}$$

The equivalent circuit is therefore *Figure 2.26*, and on reconnecting the load—of resistance R_L —we find the output current is

$$\frac{E \frac{R_2}{R_1 + R_2 + r}}{\frac{R_2(R_1 + r)}{R_1 + R_2 + r} + R_L}$$

REAL GENERATORS

and the output voltage

$$R_L \frac{E \frac{R_2}{R_1 + R_2 + r}}{\frac{R_2(R_1 + r)}{R_1 + R_2 + r} + R_L}$$

In a potentiometer $R_1 + R_2$ is constant. Let it equal R_p , then the output voltage is

$$\frac{E \cdot R_L \cdot R_2}{R_2(R_p - R_2 + r) + R_L(R_p + r)}$$

Remembering that shaft rotation is proportional to R_2 , let us plot the output voltage as a function of R_2 for various R_L for 3 types of generator: r large

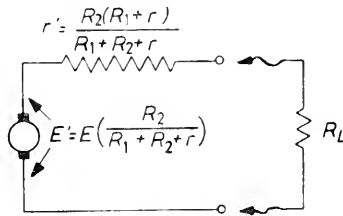


Figure 2.26

(constant-current type generator), $r = R_L$ (generator matched to load), and r small (constant-voltage type generator).

(1) For the constant-current case, the equation reduces to

$$\text{output voltage} = \frac{E \cdot R_L \cdot R_2}{r \cdot R_L + R_2}$$

(2) For the matched case,

$$\text{output voltage} = E \frac{R_2 R_L}{(R_2 + R_L)(R_p + R_L) - R_L^2}$$

(3) For the constant-voltage case,

$$\text{output voltage} = E \frac{R_2 R_L}{R_2(R_p - R_2) + R_L R_p}$$

These are plotted in *Graphs 2, 3 and 4*.

In case 3, potentiometer fed from constant-voltage type generator, we see that the relationship between volts output and shaft rotation is nearly linear when $R_L = 10 R_p$, and is moderately satisfactory when $R_L = R_p$. When $R_L = 0.1 R_p$, which is the best condition of the three in-so-far as the least power is wasted in the potentiometer, the control is downright bad; half the output voltage is controlled in 9/10ths of the shaft travel, and the other half in 1/10th.

In case 1, potentiometer fed from constant-current type generator, linearity is again optimum when $R_L = 10 R_p$, and the maximum output voltage available is 90 per cent of that obtainable when R_L is infinite, i.e. $E(R_p)/r$. At $R_L = R_p$ the linearity has deteriorated and the maximum output is down to 50 per cent. At $R_L = 0.1 R_p$ the maximum output is below 10 per cent

RESISTANCES

and the linearity so poor as for the control to be practically useless. It is worth noting that at $R_L = 0.5 R_p$ the form of the curve is approximately parabolic, which implies that if we square the ordinates we shall get a line which is substantially straight. Since the square of the ordinates is (output voltage)² and since the power delivered to a load resistance is V^2/R , it appears that we have here an arrangement for securing approximately linear control of power.

Case 2 requires a little care in interpretation. Since the load is matched to the generator, in varying the resistance of the load with respect to that of the potentiometer we are also varying the resistance of the generator. This explains why $R_L = 10 R_p$ is the lowest curve instead of the highest. It is probably better to visualize the potentiometer resistance as being varied with respect to the other two. The following then emerges: when the potentiometer resistance is small compared with the load ($R_L = 10 R_p$) the linearity of control is good but the maximum output is small. As R_p is raised the maximum output rises, till at $R_L = 0.5 R_p$ the linearity is still tolerable and the maximum output 40 per cent of the generator e.m.f. Thereafter something rather surprising happens; the maximum output continues to rise (asymptotically towards 50 per cent of E) but the output at other potentiometer settings collapses to a lower value than before. Once more, then, we find that *conditions which give low wastage of power in R_p also yield unsatisfactory control characteristics.*

There is another respect in which potentiometer control is unsatisfactory. It is not hard to see that as the potentiometer is varied, the resistance 'seen' by the generator 'looking into' the potentiometer also varies. So does the resistance seen by the load looking back into the potentiometer. Many real generators and loads, such as valve power amplifiers and penwriters, only work properly when driving a load of, or being driven by a generator of, a certain fixed resistance. In the next section we study the attenuator, which is an arrangement for controlling the transfer of power in which both generator and load 'see' a constant resistance.

The constant resistance attenuator

In *Figure 2.27* we have a generator of internal resistance r feeding a 'black-box' which in turn feeds a load R , equal to r . The black-box draws power from the generator, passes a certain fraction on to the load and

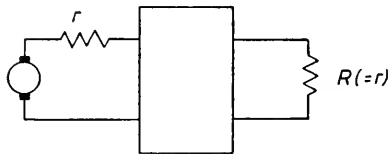


Figure 2.27

dissipates the rest within itself (as heat). This fraction is frequently adjustable, but however large or small it is the generator must always see a resistance r looking forward into the attenuator with the load connected, and the load must always see a resistance $R = r$ looking back into the attenuator with the generator connected.

REAL GENERATORS

Since the problem has left-right symmetry, it is a reasonable guess that the resistances in the black-box will have left-right symmetry too. A series resistance alone, or a parallel resistance alone, have this symmetry, but clearly they will not do. What about a combination of the two? A symmetrical arrangement worth trying is the star or *T* network, and another is the delta or pi network. Since we have seen that any delta can be replaced by an appropriate star, clearly if we can get one to do what is required, the other will do it too.

Let us then investigate the *T* network used as an attenuator. In *Figure 2.28*, if we write down the resistance looking in at the input terminals of the

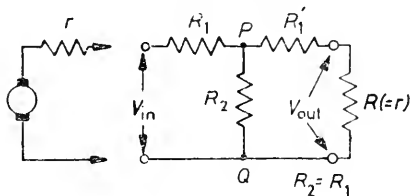


Figure 2.28

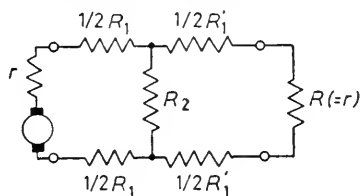


Figure 2.29

attenuator with the load connected, and equate it to the generator resistance

$$r = R_1 + \frac{R_2(R_1' + r)}{R_2 + R_1' + r}$$

the matching conditions are met. Further, if the transmission factor of the attenuator is θ , then

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \theta = \frac{V_{\text{out}}}{V_{PQ}} \cdot \frac{V_{PQ}}{V_{\text{in}}}$$

and it is not too hard to see that this equals

$$\frac{r}{R_1' + r} \cdot \frac{\frac{R_2(R_1' + r)}{R_2 + R_1' + r}}{R_1 + \frac{R_2(R_1' + r)}{R_2 + R_1' + r}}$$

In general we know r and θ and have to find R_1 and R_2 . Solving the equations we find

$$R_1 = r \frac{1 - \theta}{1 + \theta}$$

and

$$R_2 = r \frac{2\theta}{1 - \theta^2}$$

A variant of the *T* attenuator which appears at first sight more complicated than it really is, is the 'balanced' or *H* form (*Figure 2.29*). All that has been done to derive this is:

(1) to take half of R_1 and connect it in series with the other side of the generator;

(2) to take half of R_1' and connect it in series with the other side of the load.

RESISTANCES

Clearly this modification makes no difference to the currents flowing in the various branches of the circuit; but it gives the arrangement an additional axis of symmetry which is sometimes useful.

By applying the star-delta transformation to the T attenuator, we can write down the equations for the resistances in the pi form (Figure 2.30). They are:

$$R_M = R_2 + R_1 + \frac{R_1 R_2}{R_1}$$

$$R_N = R_1 + R_2 + \frac{R_1^2}{R_2}$$

Substituting the expressions found for R_1 and R_2 in terms of θ and r yields

$$R_M = r \frac{1 + \theta}{1 - \theta}$$

$$R_N = r \frac{1 - \theta^2}{2\theta}$$

A variant, related to the pi network in the same way as the H attenuator is to the T , is the box or O attenuator (Figure 2.31). This is another balanced

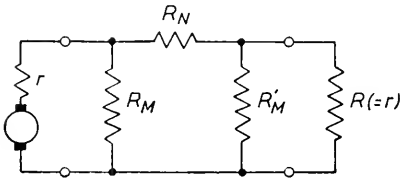


Figure 2.30

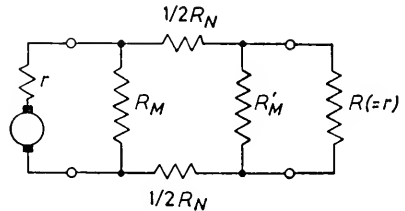


Figure 2.31

form, in which R_N is split into two halves and one placed in each horizontal limb.

Varying the attenuation—An attenuator for which θ can be varied is a useful piece of apparatus, and we consider now how this can be done. In principle θ can be varied either smoothly—continuous control—or in discrete steps. At first sight continuous control would seem to be more satisfactory, since θ can be set to any desired value; we consider it first.

Continuous control of attenuation—Clearly, to reduce θ we have—in the T attenuator, for example—to make the R_1 's larger and R_2 smaller. We could just have three calibrated variable resistances, separately adjustable, and secure any desired value of θ by working out from the equations already derived the necessary values of R_1 and R_2 , and setting them on the resistances. Such an arrangement has the merit of flexibility in that the attenuator so formed can be used over a range of generator (and load) resistances. However it would be rather clumsy in use, as to alter the θ value 3 knobs would require adjustment. The question then arises whether all three resistances can be operated from the same shaft, that is, whether they can be 'ganged'. The answer is that theoretically they can, but that if they are the resulting attenuator will have to be restricted to work between generators and loads of a

REAL GENERATORS

particular resistance. It is spoken of as, for example, a '600 ohms attenuator'. The restriction is not as serious as might at first appear, for there are certain, rather few, values of load and generator resistance which are either conventional or which occur naturally in applications where attenuators are used. The important ones are 80, 15 and 600 ohms; 15 ohms is the conventional resistance for high quality loudspeakers, 80 ohms is the apparent resistance of a dipole aerial, and 600 ohms is important in telephone engineering.

Having fixed a 'characteristic resistance' for our attenuator we can now work out the range of resistance values required for a given range of θ , and it is here that we meet a serious practical difficulty. If θ has to go down to zero the series resistance ought to increase smoothly to infinity: if θ has to reach unity the shunt resistance ought to increase smoothly to infinity. Neither can be achieved exactly or even approximated to very easily, and for this reason wide-range (i.e. $0 < \theta < 1$) constant resistance attenuators are never found. Many electronic instruments appear to possess them, but in fact they contain what might be called 'bogus attenuators', i.e. a potentiometer feeding a valve.

Suppose we restrict the attenuation range required to $0.9 < \theta < 1$, i.e. the maximum amount of attenuation our device can produce is small. This is a reasonable thing to do, since greater amounts of loss can be introduced by following our continuously variable control by a stepped one, so that the former is a 'fine' control and the latter a 'coarse'. Then the resistance values required for the *T* configuration are plotted in *Graph 5*. R_1 behaves in a manageable manner, but R_2 goes off to infinity, as expected. Suppose we leave R_2 out altogether; then our attenuator reverts to a simple variable series resistance in which R_1 and R_1' can be combined (*Figure 2.32*) into a

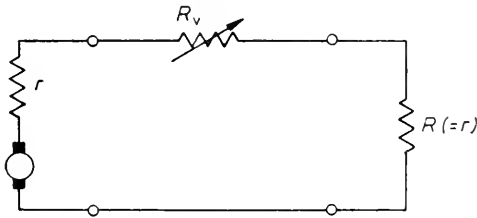


Figure 2.32

single resistance R_T . When $R_T = 0$, $\theta = 1$, and the matching is exact. When $R_T = 1/9$ th of r , $\theta = 0.9$, and the resistance seen by the generator and the load is 11 per cent high. For many purposes this would not represent a serious degree of mismatch.

In conclusion, then, the wide-range continuously variable attenuator is impossible to make. The narrow-range attenuator, however, useful in conjunction with stepped attenuators as a 'fine' control, whilst still not exactly realizable in practice, can be approximated to with sufficient accuracy by a simple variable resistance. The degree of mismatch produced is determined by the amount of fine control required.

The stepped attenuator—For wide-range variation of θ we employ the stepped attenuator in which fixed resistances are controlled by ganged

RESISTANCES

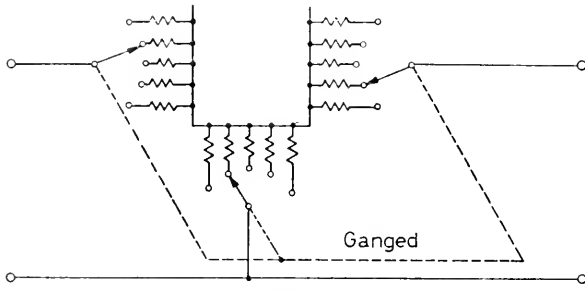


Figure 2.33

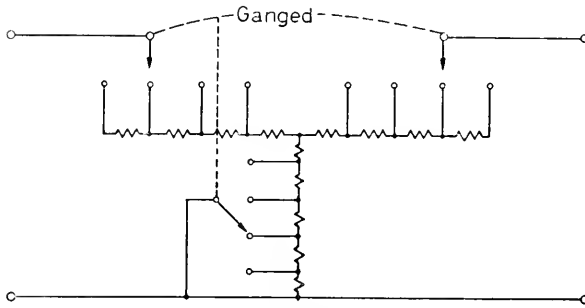


Figure 2.34

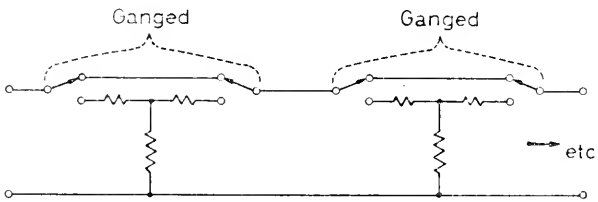


Figure 2.35

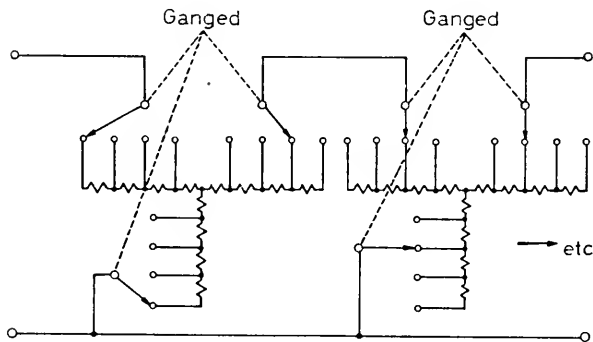


Figure 2.36

switches. We shall continue to discuss the T form, but it must be remembered that these remarks apply equally to the π configuration, also to the H and O forms, which are only trivial modifications of the T and π .

The θ of a stepped attenuator may be varied by:

- (1) changing the resistances in use for three new ones (*Figure 2.33*);
- (2) using tapped resistances, which is sometimes more economical (*Figure 2.34*);
- (3) having a varying number of fixed T 's in cascade (*Figure 2.35*). Clearly if a number of attenuator 'sections' are designed to work between a certain generator and load, i.e. have the same characteristic resistance, then they can be connected together to form a chain and their θ 's will be multiplicative.
- (4) Having a fixed number of varying T 's in cascade (*Figure 2.36*). Here again, the overall θ is the product of the θ 's for the several sections.

The decibel (abbreviation: dB)—The transmission factor, θ , was introduced because it enables us to compute the resistance values required for an attenuator section. In dealing with a number of sections, however, there is a much more convenient unit—the decibel.

Suppose we have a collection of attenuator sections, all of the same characteristic resistance, having θ values of 0.1, 0.2, 0.3 . . . 0.9, and suppose we have to make up an attenuator of overall $\theta = 0.126$. It takes a little time to see that the way to do this is to use the 0.2, 0.7 and 0.9 sections, whereas no one has any difficulty in seeing which coins from a handful of loose change go to make, say, 6/4d. It is, of course, easier to add than to find factors. Thus, if for our attenuator sections we can find a unit which expresses the input-output ratio and which is *additive* when sections are connected in cascade, then the calculations become very easy.

The Bel is defined as

$$\log_{10} \frac{P_1}{P_2}$$

where P_1 and P_2 are two powers.

Thus the Bel can be used to compare powers flowing in the same part of a circuit at different times—an important application in connection with filters—or to compare powers in two different circuits at the same time, which is the application relevant to attenuators.

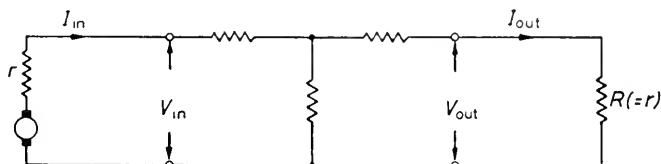


Figure 2.37

The Bel is rather a large unit, and a more popular one which avoids fractions and decimal points is the dB, which is 1/10th of a Bel. One dB is of physiological interest, in that an increase of one dB in the power supplied to a telephone receiver or loudspeaker is about the smallest increment that can be detected by the ear.

Figure 2.37 shows a generator matched to a load with a single section T

RESISTANCES

attenuator interposed. The 'attenuation' or 'insertion loss' of the section, in dB's, is

$$10 \log_{10} \frac{P_{in}}{P_{out}}$$

but $P_{in} = \frac{V_{in}^2}{r}$ and $P_{out} = \frac{V_{out}^2}{R}$

and since generator and load are matched, $r = R$, then

$$\frac{P_{in}}{P_{out}} = \frac{V_{in}^2}{V_{out}^2} = \theta^2$$

$$\therefore \text{Attenuation} = 10 \log_{10} \left(\frac{V_{in}}{V_{out}} \right)^2 = 20 \log_{10} \frac{V_{in}}{V_{out}} \text{ or } 20 \log_{10} \frac{1}{\theta} \text{ dB's}$$

Similarly it can be shown that attenuation in dB's also equals $20 \log_{10} I_{in}/I_{out}$.

Graph 6 may be used to convert a voltage or current ratio, θ , into the equivalent loss in dB's. In using it, the only point to bear in mind is that the two voltages being compared must appear across, or the currents being compared must be flowing through, equal resistances.

If a number of attenuator sections have voltage ratios $\theta_1, \theta_2, \theta_3 \dots$ and they are connected in cascade, the overall transmission factor or voltage ratio is $\theta_1\theta_2\theta_3 \dots$, and the overall attenuation in dB's is

$$20 \log_{10} \frac{1}{\theta_1\theta_2\theta_3 \dots}$$

but this is equal to

$$20 \log_{10} \frac{1}{\theta_1} + 20 \log_{10} \frac{1}{\theta_2} + 20 \log_{10} \frac{1}{\theta_3} \dots$$

that is, the attenuation of individual sections is additive.

We can now say a little more about the stepped attenuator. If only 10 or 12 degrees of attenuation are required, perhaps 36 dB's in 12 steps of 3 dB's, then types 1 or 2 are quite satisfactory; there is no difficulty in making a 12 way, 3 pole switch. If we wanted up to 100 dB's in 1 dB steps, a single 100 way switch would be quite impractical; a much more satisfactory way would be to make a 'decade' attenuator having two knobs, one giving 10 steps of 1 dB, and one giving 9 steps of 10 dB's. Decade attenuators would be of type 4.

Type 3 is called the key-switch attenuator, and uses the property of a sequence such as 1, 2, 2, 5, 10, 20, 20, 50 . . . that any number up to a maximum (given by the sum of the sequence as far as it has been taken) may be made by adding together appropriate numbers from the sequence. For example, an attenuator having 6 switches, which control respectively a 1 dB, 2 dB, 2 dB, 5 dB, 10 dB and 20 dB section, can produce an attenuation of any whole number of dB's from zero to 40.

REAL GENERATORS

The 'bridged T' attenuator—There is a useful and interesting variation of the T section, which, though it contains 4 resistances instead of 3, requires only 2 of them to be varied to vary θ . It may be derived as follows.

Take an ordinary T section (Figure 2.38); take the R_2 and split it into

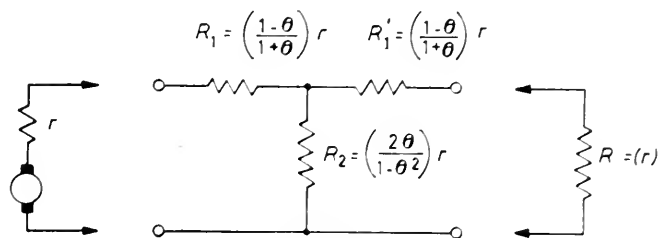


Figure 2.38

two resistances $\{\theta/(1 - \theta)\}r$ and $\{\theta/(1 + \theta)\}r$ (Figure 2.39). These add up to $(2\theta)/(1 - \theta^2)$ as before. Now transform the star between A, B, and C into a

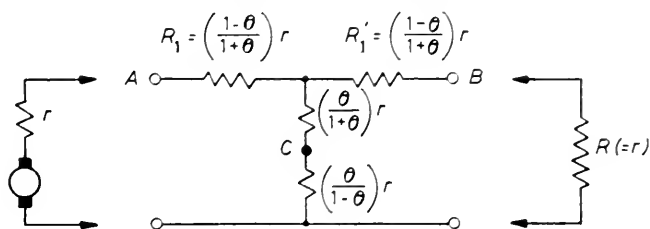


Figure 2.39

delta (Figure 2.40). This gives the 'bridged T'. Notice that two of the four resistances are fixed and equal the characteristic resistance of the section. The other two have values which are elegantly reciprocal, and easy to remember.

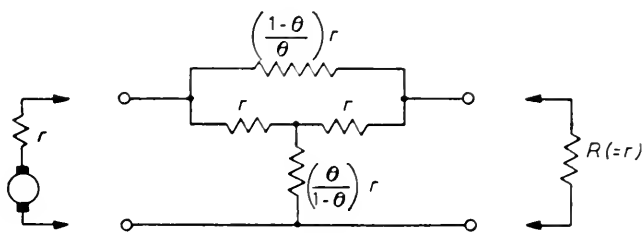


Figure 2.40

Alternating current in circuits containing only resistance—R.M.S. values

All the expressions so far derived have begun from the assumption of a direct current or direct voltage generator. They are all equally true for alternating current. Since the variable 'time' has not entered into any of the equations (except those for quantity of electricity and for energy in the very first section) they can be regarded as applying either to steady conditions or to instantaneous conditions; i.e. wherever V and I have appeared, they can be replaced by instantaneous values v and i , equal respectively to $V \sin \omega t$

RESISTANCES

and $I \sin \omega t$. This will not be true of circuits in the next chapter, where the instantaneous states of circuits will be found to depend on their past histories, and we shall have to investigate the behaviour of the networks both when connected to direct and alternating supplies.

Before ending this chapter we have to touch on what is meant by an R.M.S. value in connection with alternating currents. What does it mean to buy an electric lamp for, say, a 230 V a.c. supply, when in fact the voltage is changing the whole time? Simply this: that if the lamp were connected to a d.c. supply at 230 V, the amount of light that would be given off is the same (that is, the filament temperature would be the same) as when the lamp is connected to the 230 V a.c. supply. In other words, a 230 V a.c. supply is one which produces the same heating effect as a 230 V d.c. supply.

In *Figure 2.41* the instantaneous rate of heat production is equal to the

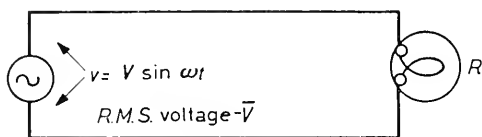


Figure 2.41

power into the lamp, which is v^2/R . Therefore the average rate of heat production is (mean of v^2/R). If this average rate could be produced by a direct voltage numerically equal to \bar{V} , then the rate of heat production would be \bar{V}^2/R and $\bar{V}^2/R = (\text{mean of } v^2/R)$

$$\begin{aligned} \text{so} \quad \bar{V} &= (\text{mean of } v^2)^{1/2} \\ &= \text{the Root of the Mean Square of } v \end{aligned}$$

For a pure sine wave generator output, $v = V \sin \omega t$, the R.M.S. value \bar{V} works out to be $0.707 V$. Note that this is not the same as the average of $V \sin \omega t$, which is 0, or even the average of $V \sin \omega t$ over half a cycle, which is $(2/\pi)V = 0.637 V$.

RESISTANCES AND CAPACITANCES

CAPACITANCES ALONE

Single capacitance

The fundamental relationship which describes the behaviour of a capacitance is as follows.

If, as a result of the effect of some generator in the circuit of which a capacitance, initially uncharged, is part, a charge Q is transferred from one plate, via the remainder of the circuit, to the other plate, and if the value of the capacitance is C , then a difference of potential V appears across its terminals, and

$$Q = VC \quad (\text{coulombs, volts, farads})$$

Capacitances in parallel

If two capacitances C_1 and C_2 are connected in parallel to a constant direct-voltage generator of e.m.f. E (Figure 3.1) then a charge Q_1 is displaced round

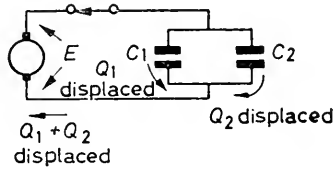


Figure 3.1

the circuit in virtue of C_1 , and a charge Q_2 in virtue of C_2 .

The effective capacitance C_{eff} for the combination is

$$\begin{aligned} & \frac{\text{Total charge displaced}}{E} \\ &= \frac{Q_1 + Q_2}{E} = \frac{C_1 E_1 + C_2 E_2}{E} \\ &= C_1 + C_2 \end{aligned}$$

and in general, for n capacitances in parallel, $C_{\text{eff}} = C_1 + C_2 + C_3 \dots + C_n$.

If we differentiate the fundamental equation $Q = VC$ with respect to time

$$\frac{dQ}{dt} = \frac{dV}{dt} C$$

but $\frac{dQ}{dt} = I, \quad \therefore I = \frac{dV}{dt} C$

Thus if a capacitance, initially uncharged, be connected to a direct constant-current generator of output I , the potential difference across the capacitance grows linearly with time, at a rate $dV/dt = I/C$ (Figure 3.2).

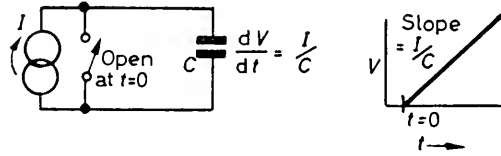


Figure 3.2

Capacitances in series

If two capacitances C_1 and C_2 are connected in series with a generator of constant current I , the rate of growth of potential difference across C_1 is $(dV_{C_1})/(dt) = I/C_1$ and across C_2 is $(dV_{C_2})/(dt) = I/C_2$ (Figure 3.3).

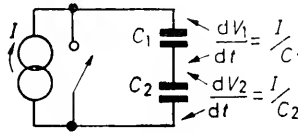


Figure 3.3

Therefore the rate of growth of potential difference across the combination is

$$\frac{I}{C_1} + \frac{I}{C_2} = I \left(\frac{1}{C_1} + \frac{1}{C_2} \right)$$

If the effective capacitance of C_1 and C_2 in series is C_{eff} , the rate of growth of potential difference across C_{eff} when charged by a current I is

$$\frac{I}{C_{\text{eff}}}$$

therefore $\frac{I}{C_{\text{eff}}} = I \left(\frac{1}{C_1} + \frac{1}{C_2} \right)$ so $\frac{1}{C_{\text{eff}}} = \frac{1}{C_1} + \frac{1}{C_2}$

and in general for n capacitances in series is

$$\frac{1}{C_{\text{eff}}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} \dots + \frac{1}{C_n}$$

A more convenient form for two capacitances in series is $C_{\text{eff}} = (C_1 C_2) / (C_1 + C_2)$.

Capacitance connected to constant-voltage alternator

If a capacitance be connected to an alternating voltage, constant-voltage generator, or constant-voltage alternator (Figure 3.4), of output $v = V \sin \omega t$, the current through the capacitance is

$$\begin{aligned} i &= C \, dv/dt \\ &= C \frac{d(V \sin \omega t)}{dt} \\ &= \omega C V \cos \omega t \end{aligned}$$

RESISTANCE AND CAPACITANCE IN SERIES

The voltage and current waveforms are sketched in *Figure 3.5*. Both have the same form, but not the same phase, and in fact *the current through a capacitance leads the terminal voltage by 90 degrees in an alternating current circuit.*

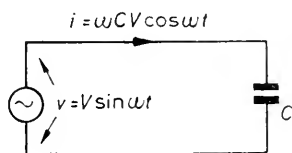


Figure 3.4

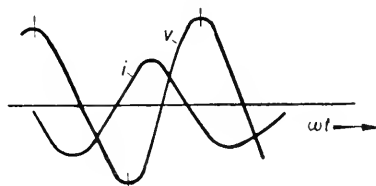


Figure 3.5

Reactance

The value of a resistance in ohms is given by the ratio of potential difference to current; the analogous property of a capacitance is also measured in ohms and is called the *capacitive reactance*, symbol X_C . Numerically it equals $V/(\omega CV) = 1/(\omega C)$, the phase difference represented by the sine and cosine terms being ignored. Capacitive reactance is conventionally regarded as being negative in sign, for a reason which appears in Chapter 5.

Power in capacitances

In *Figure 3.4* the instantaneous power supplied to the capacitance is

$$\begin{aligned} P &= vi \\ &= V \sin \omega t \cdot \omega CV \cos \omega t \\ &= V^2 C \omega \sin \omega t \cdot \cos \omega t \\ &= \frac{1}{2} V^2 C \omega \sin 2\omega t \end{aligned}$$

This function moves positive and negative symmetrically, which means that power surges back and forth into the capacitance and out again. On the average the power supplied is

$$\frac{1}{2} V^2 C \omega \cdot \frac{1}{2\pi} \int_0^{2\pi} \sin 2\omega t \, dt = 0$$

Therefore, no power is consumed by a capacitance.

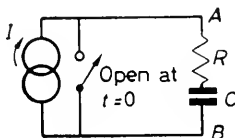


Figure 3.6

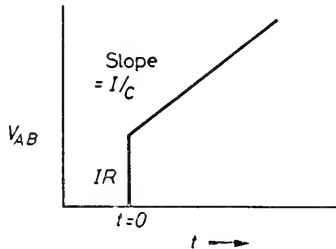
RESISTANCE AND CAPACITANCE IN SERIES

Series R and C connected to a constant direct current generator

If a resistance and a capacitance are connected in series to a generator of constant direct current I (*Figure 3.6*), upon opening the switch a waveform is

RESISTANCES AND CAPACITANCES

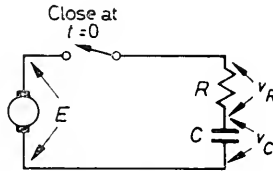
produced across $A-B$ of the form of *Figure 3.7*; there is an initial step of magnitude IR , followed by a steady rise in voltage of rate I/C . This follows straightforwardly from our findings for a resistance and a capacitance connected to a constant-current generator separately.



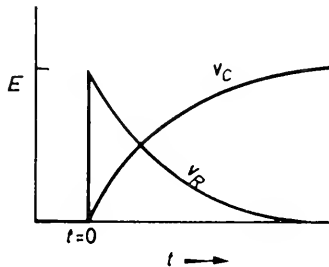
Series R and C connected to a constant direct voltage generator

In this case the solution is not quite so simple. In *Figure 3.8*, if the generator is of e.m.f. E , then clearly

$$\begin{aligned} E &= v_R + v_C \\ &= iR + \int_0^t dV \\ &= iR + \int_0^t \frac{i}{C} dt \end{aligned}$$



This is a differential equation having the solution, $i = (E/R)e^{-t/RC}$. The voltage across the resistance, v_R , is $iR = Ee^{-t/RC}$, and the voltage across the capacitance, v_C , must be E minus this, $E(1 - e^{-t/RC})$ (*Figure 3.9*).



For the non-mathematically minded reader, a verbal description of what is happening is as follows. When the generator is first connected, the capacitance is uncharged, and has no potential difference across it; therefore the

RESISTANCE AND CAPACITANCE IN SERIES

whole of the e.m.f. appears across the resistance, causing a current E/R to flow through it to charge the capacitance, and the capacitance voltage begins to rise at a rate $dv/dt = i/C = E/CR$. However, as soon as the voltage across the capacitance rises, that across the resistance must fall, since their sum must equal E : therefore the current falls, so that the capacitance voltage may continue to rise, but more and more slowly. The final state of affairs, which theoretically takes an infinite time to reach, is that the capacitance voltage just reaches the generator voltage as the current falls to zero, when circuit action ceases.

This circuit is one of an extremely important class possessing an 'exponential response' which occurs frequently in electronics, and it is important to be thoroughly at home with it.

An important notion associated with the exponential response is that of 'time constant', which is measured in seconds and is a measure of the rapidity with which the circuit responds to the application of E . Clearly we cannot use the time taken for circuit action to cease, since this is theoretically infinite. Time constant is defined simply as

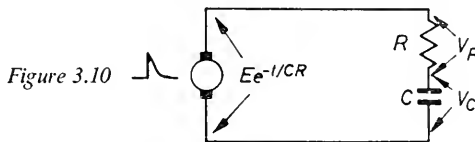
$$T = RC \quad (\text{seconds, ohms, farads, or more usefully seconds, megohms microfarads})$$

Its meaning may be seen in two ways:

(1) From the verbal description. The capacitance charges at an initial rate E/CR . If this rate could be maintained, the time taken to reach the charging voltage E would be $E/(E/CR) = CR$. Thus *the time constant is the time it would take the circuit to reach equilibrium if the initial capacitance charging rate could be maintained.*

(2) From the mathematical description. The capacitance voltage is $E(1 - e^{-t/RC})$. When t reaches $T = CR$, the time constant, then $V_C = E(1 - e^{-1}) = 0.632 E$, that is, *the time constant is the time taken by the capacitance voltage to reach 63.2 per cent of its final value.*

The two pieces of analysis which follow are important in connection with multi-stage capacitance coupled amplifiers, such as are often employed in electrophysiology; their importance will become evident later.



Resistance and capacitance in series with a constant-voltage generator whose output is of the form $e = Ee^{-t/CR}$.

Here we are feeding our series circuit with a voltage waveform similar to v_R in Figure 3.9 (Figure 3.10). This time we have

$$\begin{aligned} Ee^{-t/CR} &= v_R + v_C \\ &= iR + \int_0^t dv \\ &= iR + \int_0^t \frac{i}{C} dt \end{aligned}$$

RESISTANCES AND CAPACITANCES

The solution is

$$i = \frac{E}{R} e^{-t/CR} \left(1 - \frac{t}{CR} \right)$$

giving for the voltage across the resistance

$$v_R = E \cdot e^{-t/CR} \left(1 - \frac{t}{CR} \right)$$

and for the voltage across the capacitance,

$$v_C = E \cdot e^{-t/CR} \frac{t}{CR}$$

Resistance and capacitance in series with a generator whose output is of the form $e = Ee^{-t/CR}(1 - t/CR)$

We repeat the above process once more, supplying our series circuit with

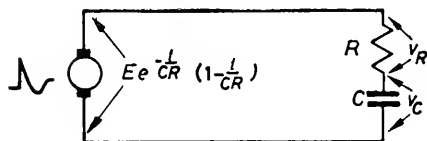


Figure 3.11

a voltage waveform similar to v_R in Figure 3.10 (Figure 3.11). By a similar procedure we find

$$v_R = E \cdot e^{-t/CR} \left(1 - \frac{2t}{CR} + \frac{t^2}{2C^2R^2} \right)$$

and

$$v_C = E \cdot e^{-t/CR} \left(\frac{t}{CR} + \frac{t^2}{2C^2R^2} \right)$$

Graph 7 is a plot of v_R in the three cases, i.e.

Case 1. Voltage step after passage through one RC network.

$$E \cdot e^{-t/CR}$$

Case 2. Voltage step after passage through two RC networks.

$$E \cdot e^{-t/CR} \left(1 - \frac{t}{CR} \right)$$

Case 3. Voltage step after passage through three RC networks

$$E \cdot e^{-t/CR} \left(1 - \frac{2t}{CR} + \frac{t^2}{2C^2R^2} \right)$$

Notice that the response in cases 2 and 3 is oscillatory.

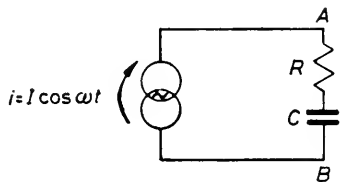
Resistance and capacitance in series with constant-current alternator

When a resistance and a capacitance are connected in series to a generator

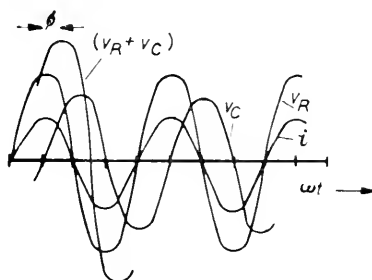
RESISTANCE AND CAPACITANCE IN SERIES

of constant current $i = I \cos \omega t$ (Figure 3.12) we have $v_{AB} = v_R + v_C$. Now $v_R = IR \cos \omega t$ and we have seen that $v_C = (I/\omega C) \cdot \sin \omega t$.

The question is how to add these to find v_{AB} . We could do it by plotting



out v_R and v_C and adding the ordinates for a number of values of ωt (Figure 3.13). We should then find v_{AB} to be another wave of sinusoidal form and

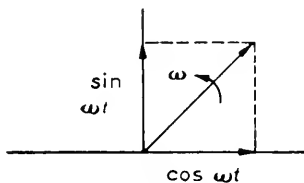


of phase intermediate between v_R and v_C . If it is found to lag by an angle ϕ behind the curves for i and v_R , then it may be expressed as

$$v_{AB} = A \cos(\omega t + \phi)$$

ϕ is called the phase angle for the network and is the angle between the waves of terminal voltage and current; it lies between 0 and 90 degrees. A is called the modulus of v_{AB} . The modulus of a wave refers to its amplitude, and says nothing about its phase; it may be written v_{AB} , which is read as 'mod v_{AB} '.

There is another, more sophisticated, approach which gets the answer a good deal quicker. The quantities $\sin \omega t$ and $\cos \omega t$ may be regarded as being generated by the projections of a unit vector, rotating in the plane of



the paper at ω radians per second, on to vertical and horizontal axes (Figure 3.14). Then $\cos \omega t$ and $\sin \omega t$ may themselves be regarded as vectors, associated with perpendicular directions. Thus when $\cos \omega t$ is multiplied

by the scalar IR , the product is a directed quantity (Figure 3.15). Similarly when $\sin \omega t$ is multiplied by the scalar $-I/\omega C$ the product is another directed quantity (downward, because capacitive reactance is by convention negative) (Figure 3.16). Their sum is obtained by vector addition (Figure 3.17). The

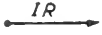


Figure 3.15

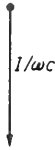


Figure 3.16

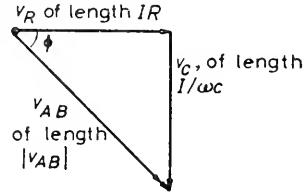


Figure 3.17

angle between v_R and v_{AB} is the phase angle ϕ , which is now seen to be equal to

$$\tan^{-1} \frac{I/\omega C}{IR} = \tan^{-1} \frac{1}{\omega CR}$$

and the length of v_{AB} , which gives us $|v_{AB}|$, is now seen by Pythagoras to be

$$\left\{ (IR)^2 + \left(\frac{I}{\omega C} \right)^2 \right\}^{1/2} = I \left\{ R^2 + \left(\frac{1}{\omega C} \right)^2 \right\}^{1/2}$$

Thus the complete description of v_{AB} is

$$v_{AB} = I \overbrace{\left\{ R^2 + \left(\frac{1}{\omega C} \right)^2 \right\}^{1/2}}^{\text{modulus}} \underbrace{\cos(\omega t + \phi)}_{\substack{\text{phase angle} \\ \downarrow}} \quad \text{where } \phi = \tan^{-1} \frac{1}{\omega CR}$$

Impedance—The impedance of the circuit is the ratio of terminal voltage to current, is measured in ohms, and is symbolized by Z . It expresses with circuits containing resistances and capacitances the amount of opposition to the flow of current, in a manner analogous to the notions of resistance and reactance in circuits containing, respectively, resistors and capacitors only. The presence of the phase difference between voltage and current makes impedance a ‘complex number’, containing a modulus part and an angle part. If, for the present, we restrict ourselves to the modulus, then

$$|v_{AB}| = I|Z| \quad (\text{cf. } V = IR)$$

and on comparing this with the expression for v_{AB} above

$$|Z| = \left\{ R^2 + \left(\frac{1}{\omega C} \right)^2 \right\}^{1/2}$$

or, remembering that capacitive reactance, $X_C = 1/\omega C$

$$|Z| = \{R^2 + (X_C)^2\}^{1/2}$$

The network we have been discussing is an extremely simple one, yet the analysis has yielded a square root term, and square root terms are notoriously tiresome in calculations. It is not hard to see that in the analysis of more elaborate resistance and capacitance networks, the equations are liable to become very unwieldy. There is an analytical technique which we shall use

RESISTANCE AND CAPACITANCE IN SERIES

from now on which greatly facilitates the work. It is the use of the j operator.

The j operator—The expression for the impedance of a resistance and a capacitance in series is $\{R^2 + X_C^2\}^{1/2}$ and to get a numerical answer the processes of squaring, adding, and square-rooting must at some time be carried out. However, in solving networks containing many R 's and C 's, it is possible to delay the appearance of square root terms until the very end of the calculation, and only to have to work them out once. The procedure is to label all the reactances, from the outset, with the letter j , which signifies that they are impedances of a different kind from the resistances. Thereafter the difference can be forgotten, and the theorems advanced for series and parallel connections, Thévenin's theorem, and the star-delta transformation, may be extended to cover the solution of networks containing both resistances and capacitances mixed. Whatever the nature of the problem being solved, one emerges from the calculations with an expression of the form: $A + jB$, and only then is it necessary to square, add, and root to get the numerical answer, $\{A^2 + B^2\}^{1/2}$.

In *Figure 3.18* we see a scale of numbers, and on it a vector which represents

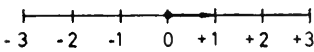


Figure 3.18

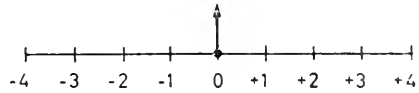


Figure 3.19

$+1$. If j operates on $+1$, (symbolically, $j(+1)$) it has the effect of rotating the vector anti-clockwise through an angle of 90 degrees (*Figure 3.19*). If we allow j to operate again, the vector undergoes a further 90 degrees rotation (*Figure 3.20*). Symbolically, *Figure 3.20* is $j\{j(+1)\}$. j obeys the ordinary rules of algebra, so that $j\{j(+1)\} = j^2(+1)$ or simply j^2 .

Looking at *Figure 3.20*, it is clear that the vector also equals -1 , therefore $j^2 = -1$ and $j = (-1)^{1/2}$.

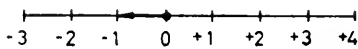


Figure 3.20

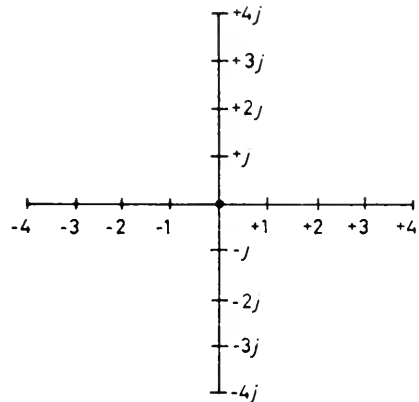


Figure 3.21

$(-1)^{1/2}$ is an 'imaginary number', and the vertically directed vector in *Figure 3.19* is marking off one division in the 'scale of imaginary numbers'. Filling in this scale, we get *Figure 3.21*.

If j is allowed to operate a few more times, it becomes evident that:

$$j^3 = -j$$

$$j^4 = +1$$

and that

$$j = -\frac{1}{j}$$

Now, when we were considering the series C and R connected to the constant-current alternator, in order to find the potential difference across the combination we had to add the voltage across the resistance and the voltage across the capacitance vectorially, 'at right angles'. With the help of j we can deal with this easily, for by convention resistance is regarded as being measured along the axis of real positive numbers, and capacitive reactance along the axis of *negative* imaginary numbers; thus the expression for the impedance is just $R - jX_C$. Let us now apply j to solving a practical problem.

Series R and C connected to a constant-voltage alternator

With the arrangement of *Figure 3.22*, if the generator output is represented by

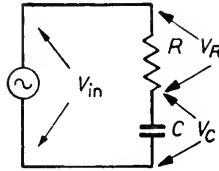


Figure 3.22

vector V_{in} the current which flows round the circuit is represented by the vector

$$I = \frac{V_{in}}{R - jX_C}$$

and the potential difference it produces across R is

$$V_R = \frac{V_{in}R}{R - jX_C}$$

and across C is

$$V_C = \frac{-jV_{in}X_C}{R - jX_C}$$

Thus

$$\frac{V_C}{V_{in}} = \frac{-jX_C}{R - jX_C}$$

and

$$\frac{V_R}{V_{in}} = \frac{R}{R - jX_C}$$

Notice that these expressions may be written down by inspection; for the network forms a potentiometer which is only different from the potentiometers we have so far dealt with in that one of the elements is reactive

instead of resistive. The expression for the transmission factor is similar to that for the all-resistance potentiometer, but it contains some terms labelled by $-j$.

The resistance-capacitance high-pass filter—With the arrangement of Figure 3.23, the transmission factor

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{R}{R - jX_C} \quad \text{Now } X_C = \frac{1}{\omega C} \quad \therefore \frac{V_{\text{out}}}{V_{\text{in}}} = \frac{R}{R - \frac{j}{\omega C}} = \frac{1}{1 - \frac{j}{\omega CR}}$$

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{\left\{ 1 + \left(\frac{1}{\omega CR} \right)^2 \right\}^{1/2}}$$

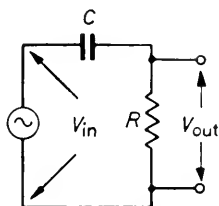


Figure 3.23

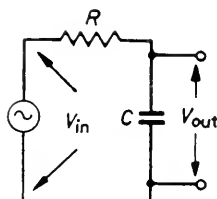


Figure 3.24

The resistance-capacitance low-pass filter—This is the arrangement of Figure 3.24. By a similar process to the above

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{\{ 1 + (\omega CR)^2 \}^{1/2}}$$

The moduli for the transmission factors for these two filters are plotted as functions of ω on linear scales in Graph 8, which effectively conceals the essential symmetry in the behaviour of the two circuits. Replotting on scales of log frequency and log [transmission factor] (i.e. a linear scale in dB's) presents the much more satisfactory picture of Graph 9. From Graph 9 it is clear why the devices are called 'filters'. In the high-pass case, for example, evidently all frequencies much above $\omega = 1/CR$ are passed without significant attenuation, whereas all those much below $1/CR$ are reduced. The frequency $1/CR$, which marks the transition between the two regimes, is called the 'turn-over frequency'; it is also sometimes, rather optimistically, called the 'cut-off frequency'. In general we shall call a filter any network whose transmission factor is frequency dependent.

Use of log frequency and linear dB scales—The use of logarithmic scales for frequency and transmission factor (the latter is, of course, the same as saying the scale is linear in dB's) is practically universal in plotting the performance of all kinds of filter, for by this means the filter properties can be most clearly exhibited. In addition, the use of these scales allows the performance of many types of filter to be sketched freehand. For example, in the case of the R-C high-pass filter, the transmission characteristic may be sketched as follows:

Mark a point, A , at $\omega = 1/CR$, insertion loss = zero.

From A , draw a dotted straight line horizontally to the right.

RESISTANCES AND CAPACITANCES

From *A*, draw a dotted straight line to the left and downwards, of slope 6* dB per octave, that is, 6 dB's for a twofold change in frequency.

From *A*, mark the point *B*, 3 dB's below *A*.

Draw a smooth curve, through *B*, asymptotic to the dotted lines.

This is the required characteristic curve (Figure 3.25). The transmission curve

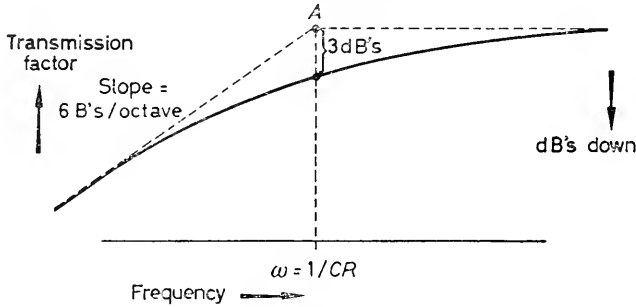


Figure 3.25

for the low-pass case is, of course, the mirror image of this and the technique for sketching it will be obvious.

*Finding the phase shift when using *j* operator*—To plot the modulus of the transmission factor for a filter is not to tell the whole story, for we are saying nothing about the relative phase angle between input and output. To find the 'phase shift' when using *j* operator we employ a process called 'rationalization'. In general the final expression for V_{out}/V_{in} has the form $(A + jB)/(C + jD)$

To rationalize this, we multiply numerator and denominator by $C - jD$, giving

$$\frac{V_{out}}{V_{in}} = \frac{(A + jB)(C - jD)}{(C + jD)(C - jD)} = \frac{(AC + BC) + j(BC - AD)}{C^2 + D^2}$$

The tangent of the phase shift is now given by the *j* part of the numerator divided by the real part, i.e.

$$\phi = \tan^{-1} \frac{BC - AD}{AC + BC}$$

If the phase of the output leads on that of the input, the phase angle will come out positive; if it lags, negative.

Thus, for the high-pass filter, we had

$$\frac{V_{out}}{V_{in}} = \frac{1}{1 - \frac{j}{\omega CR}}$$

rationalizing,

$$\frac{V_{out}}{V_{in}} = \frac{1 + \frac{j}{\omega CR}}{\left(1 - \frac{j}{\omega CR}\right)\left(1 + \frac{j}{\omega CR}\right)} = \frac{1 + \frac{j}{\omega CR}}{1 + \left(\frac{1}{\omega CR}\right)^2}$$

* 6 dB's loss is almost exactly a transmission factor of $\frac{1}{2}$ (Graph 6).

RESISTANCE AND CAPACITANCE IN SERIES

therefore the phase shift is $\tan^{-1} 1/\omega CR$. Similarly, for the low-pass filter, we had

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{-\frac{j}{\omega C}}{R - \frac{j}{\omega C}} = \frac{-\frac{j}{\omega CR}}{1 + \frac{j}{\omega CR}}$$

rationalizing

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{-\frac{j}{\omega CR} \left(1 + \frac{j}{\omega CR}\right)}{\left(1 - \frac{j}{\omega CR}\right) \left(1 + \frac{j}{\omega CR}\right)} = \frac{\frac{1}{\omega^2 C^2 R^2} - j \left(\frac{1}{\omega CR}\right)}{1 + \frac{1}{\omega^2 C^2 R^2}}$$

therefore the phase shift is

$$\tan^{-1} \frac{1}{\frac{\omega CR}{1}} = \tan^{-1} \frac{1}{\omega CR} \quad (\text{Graph 10})$$

Loaded R-C high-pass filter

If a load be connected to a simple resistance-capacitance high-pass filter (Figure 3.26) we have a new effective resistance-capacitance product $R'C$,

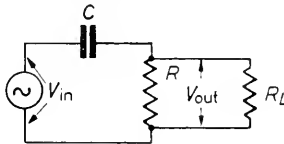


Figure 3.26

where $R' = R_L R / (R_L + R)$. Thus, the effect on the transmission characteristic of the connection of the load is to move the curve bodily to the right, to a new turn-over frequency, higher than the old one by the factor

$$\frac{R + R_L}{R_L}$$

Loaded R-C low-pass filter

If a load be connected to a simple low-pass filter (Figure 3.27), applying

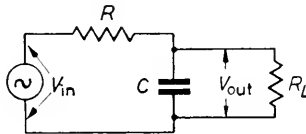


Figure 3.27

Thévenin's theorem at the capacitance terminals,
 the open-circuit e.m.f. is $V_{\text{in}} R_L / (R + R_L)$,
 the resistance looking in at the capacitance terminals is $RR_L / (R + R_L)$

RESISTANCES AND CAPACITANCES

and the equivalent circuit is thus (Figure 3.28) an unloaded filter. The transmission characteristic is again moved to the right such that the turn-over frequency is raised by a factor $(R + R_L)/R$, and it is also moved downward an amount $20 \log_{10} (R + R_L)/R$ dB's.

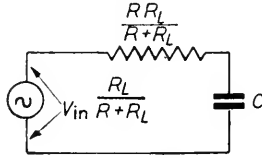


Figure 3.28

Low-pass R-C filter fed from a real generator

When a low-pass filter is fed from a real generator of internal resistance r , the effect is to alter the value of resistance required in the filter. For a given turn-over frequency ω_C , the total R required (Figure 3.29) is $1/C\omega$. Since

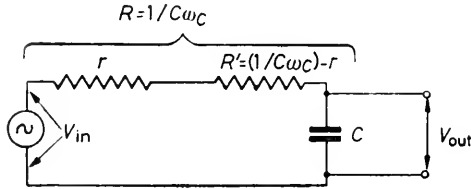


Figure 3.29

we already have r , the value of R' required is only $(1/C\omega) - r$. There is no reason why we should not dispense with R' altogether, connecting C directly across the terminals of the generator. Then the value of C required is simply $1/\omega_C r$ (Figure 3.30).

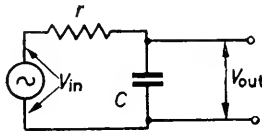


Figure 3.30

High-pass filter fed from real generator

This case (Figure 3.31) is a little more complicated. As before, the value

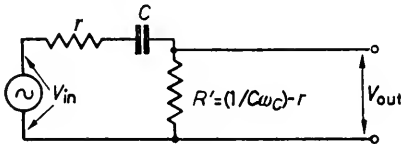


Figure 3.31

of R' required is $(1/C\omega) - r$. The output appears across both R' and r , but only the proportion across R' is usable. The filter therefore introduces—in addition to the frequency dependent attenuation—a fixed attenuation— $20 \log_{10} (R' + r)/R'$ dB's.

RESISTANCE AND CAPACITANCE IN SERIES

At this point the reader may accuse me of doing precisely what I have warned against; namely, of converting voltage ratios into decibels with insufficient reference to the resistances across which the voltages appear. He may further ask, what is the point of considering the behaviour of networks to which no useful load is attached? The answer is that the filters in *Figures 3.29* and *3.31* are frequently used to feed a valve, which constitutes a load, but a load whose resistance is substantially infinite, and which therefore does not influence the operation of the filters. If the load resistance is infinite in all cases, we can convert voltage ratios into decibels without hesitation.

As an exercise, the reader may care to consider what happens when the simple R-C high- and low-pass filters are fed from a real generator and work into a load which is not of infinite resistance.

R-C filters comprising cascaded sections

Suppose V_{in} is the sum of two alternating voltages of frequency ω_1 and ω_2 , and that we wish to filter out as much as possible of the contribution of the generator working at ω_2 , the higher frequency. We use, of course, a low-pass filter (*Figure 3.32*). If ω_1 and ω_2 are close together, the difference in the

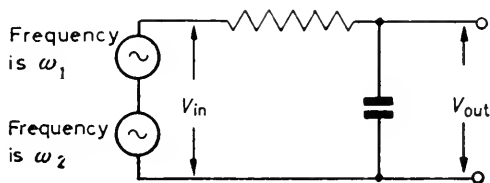


Figure 3.32

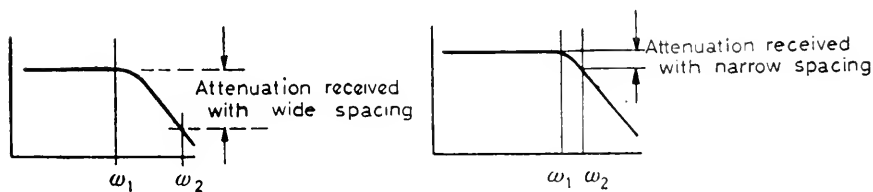


Figure 3.33

attenuations received by the two frequencies is clearly less than if they are widely separated (*Figure 3.33*). It often happens that the discrimination offered by a single R-C section is insufficient, and the question arises as to how to obtain more.

In electronics, if the output of one device is fed to the input of a similar device, the devices are said to be 'in cascade'. If two sections of R-C filtering are connected in cascade, the best that one can hope for is that the attenuation obtained from the pair of sections will be twice that obtainable from one (*Graph 11*). In fact this is not possible, owing to the nature of the loading placed upon the first section by the second.

RESISTANCES AND CAPACITANCES

In *Figure 3.34*, if $R_1 = R = R_2$ and $C_1 = C_2 = C$, applying Thévenin's theorem at the points $A-B$ and looking to the left, we see

an open-circuit voltage $\frac{-jX_C}{R - jX_C} V_{in}$

in series with an impedance $\frac{R \cdot -jX_C}{R - jX_C}$

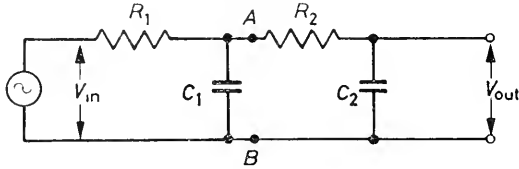


Figure 3.34

The equivalent circuit is therefore *Figure 3.35*—an elaboration of a potentiometer circuit. V_{out} is therefore

$$\left(\frac{-jX_C}{R - jX_C} V_{in} \right) \cdot \frac{-jX_C}{R - jX_C + \frac{R \cdot -jX_C}{R - jX_C}} = V_{in} \frac{X_C^2}{R^2 + X_C^2 - j(3X_C R)}$$

$$\therefore \left| \frac{V_{out}}{V_{in}} \right| = \frac{X_C^2}{\{(R^2 + X_C^2)^2 + 9X_C^2 R^2\}^{1/2}} \quad \text{Substituting } \frac{1}{\omega C} \text{ for } X_C$$

$$\left| \frac{V_{out}}{V_{in}} \right| = \frac{1}{\{(1 + \omega^2 C^2 R^2)^2 + 9\omega^2 C^2 R^2\}^{1/2}}$$

This function is also plotted in *Graph 11*. Clearly the degradation of the performance due to cascading the sections takes the form of a lack of 'squareness' in the characteristic in the region of the turn-over frequency.

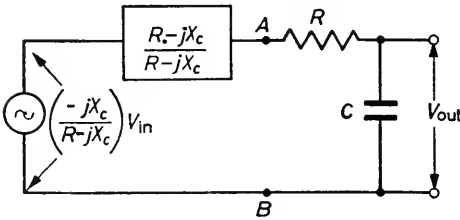


Figure 3.35

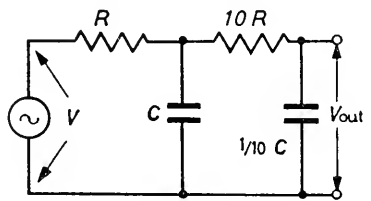


Figure 3.36

Tapered sections—The effect can be mitigated by 'tapering' the sections, that is, by making the resistance greater, and the capacitance lower, in the second section than in the first. Thus if $R_2 = 10 R_1$, and $C_2 = 1/10$ th of C_1 (*Figure 3.36*), the turn-over frequency of both sections is still the same, but the loading effect of the second on the first is much less severe; the performance will much more nearly approach the ideal for two sections, as in *Graph 11*.

Resistance and capacitance in parallel, connected to constant direct current generator

This is another circuit showing an exponential response (*Figure 3.37*). At

RESISTANCE AND CAPACITANCE IN SERIES

any time t after opening the switch

$$v_R = v_C$$

$$\therefore i_R \cdot R = \int_0^t dv_C = \int_0^t \frac{i_C}{C} dt. \quad \text{But } i_R = I - i_C$$

$$\therefore R(I - i_C) = \int_0^t \frac{i_C}{C} dt$$

The solution is $v_R = v_C = IR(1 - e^{-t/RC})$.

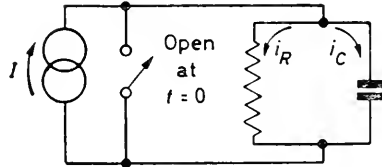


Figure 3.37

If the generator and switch are then disconnected, the capacitance discharges again, through the resistance, and $v_R = v_C = Ve^{-t/RC}$, where V is the potential difference to which the capacitance was charged. If the original charging process was substantially completed, V is, of course, equal to IR .

Resistance and capacitance in parallel, connected to constant-voltage alternator (Figure 3.38)

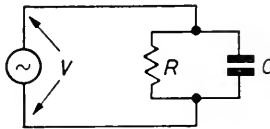


Figure 3.38

In this case, we have:

$$\text{Current in the resistance} = \frac{V}{R}$$

$$\text{Current in capacitance} = \frac{V}{-jX_C} = \frac{jV}{X_C}$$

$$\text{Therefore, total current} = V \left(\frac{1}{R} + \frac{j}{X_C} \right)$$

$$\text{So the impedance} = \frac{V}{V \left(\frac{1}{R} + \frac{j}{X_C} \right)} = \frac{1}{\frac{1}{R} + \frac{j}{X_C}}$$

Compare this with the expression for two resistances in parallel.

If we take the expression for the impedance of the parallel combination and rewrite it in the form

$$Z = \frac{-jX_C \cdot R}{R - jX_C}$$

then, rationalizing,

$$Z = \frac{-jX_C R(R + jX_C)}{R^2 + X_C^2} = \frac{X_C^2 R - jR^2 X_C}{R^2 + X_C^2}$$

This, however, is the impedance of a series combination, comprising a resistance of $(X_C^2 R)/(R^2 + X_C^2)$ and a capacitance of reactance $(R^2 X_C)/(R^2 + X_C^2)$. The phase angle is thus seen to be

$$\tan^{-1} \frac{R^2 X_C}{X_C^2 R} = \tan^{-1} \frac{R}{X_C}$$

Summarizing, then, the two networks in *Figure 3.39* are equivalent at one

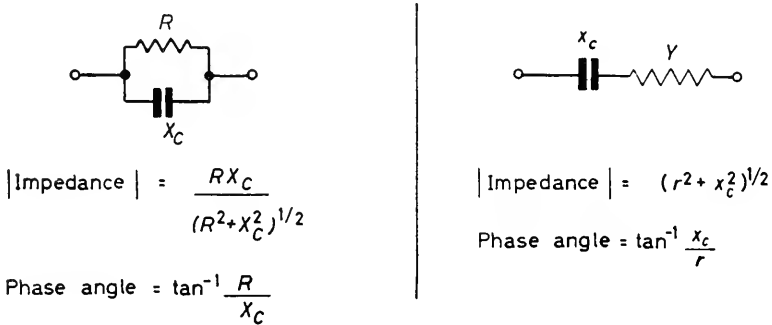


Figure 3.39

frequency, that at which

$$x_c = \frac{R^2 X_C}{R^2 + X_C^2} \quad \text{and} \quad r = \frac{X_C^2 R}{R^2 + X_C^2}$$

Potentiometer feeding a piece of screened cable

R-C filters are not always built into apparatus by design; they sometimes occur unintentionally. For example, in *Figure 3.40*, a piece of apparatus *A*

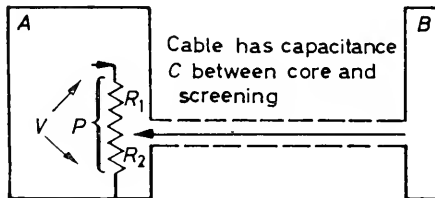


Figure 3.40

is supposed to be feeding another piece of apparatus *B* via a length of screened cable. *P* is an amplitude control, a potentiometer to adjust the quantity of signal transferred from *A* to *B*; the load imposed on *A* by *B* is assumed to be negligible.

RESISTANCE AND CAPACITANCE IN SERIES

The equivalent circuit of the arrangement is *Figure 3.41*, from which it is clear that we have a low-pass filter whose turn-over frequency depends on the amplitude control setting; at maximum output, $R_1 = 0$, so the effective

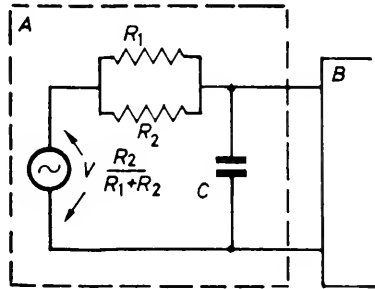


Figure 3.41

resistance in the filter is also 0, and the cable capacitance has no effect. When the slider of P is at mid-travel the effect can be serious: thus, if $P = 1$ megohm, and the cable is 6 ft. long and is P.V.C. insulated, its capacitance may well be 400 pF. At mid-setting of P , $R_1 R_2 / (R_1 + R_2)$ is maximal at 0.25 MΩ, and the turn-over frequency is

$$\begin{aligned} \omega &= 1/CR \\ &= \frac{1}{400 \times 10^{-12} \times 0.25 \times 10^6} \\ &= 10^4 \text{ radians/sec} \end{aligned}$$

Dividing by 2π , we find this to be a frequency of the order of 1.5 kc/s, much too low for many applications. Thus if A were an electrophysiological pre-amplifier, and B the main amplifier, manipulation of the amplitude control would be found to affect the shape, as well as the size, of action potentials, due to attenuation of high-frequency components in the waveform at reduced amplitude settings. Notice, however, that as the control is moved to below $\frac{1}{2}$ maximum, the position improves again.

R-C circuits which filter with limited phase shift

Inspection of *Graphs 9* and *10* reveals that the simple R-C filters, by the time they are offering an attenuation of 20 dB's, are also introducing a phase

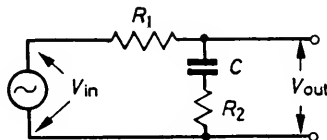


Figure 3.42

shift of nearly 90 degrees. In the chapter on feedback we shall see that there are times when we require circuits which will filter, whilst holding the phase shift within bounds.

The low-pass version of such a filter is shown in *Figure 3.42*; it is derived

RESISTANCES AND CAPACITANCES

from the simple low-pass filter by the addition of an extra resistance. By inspection

$$\frac{V_{out}}{V_{in}} = \frac{R_2 + \frac{1}{j\omega C}}{R_1 + R_2 + \frac{1}{j\omega C}}$$

We can simplify the calculation if we make an assumption which is usually permissible, that $R_2 \ll R_1$. Then, multiplying through by $j\omega C$,

$$\frac{V_{out}}{V_{in}} = \frac{1 + j\omega CR_2}{1 + j\omega CR_1}$$

Now take R' , the geometric mean of R_1 and R_2 , and define a such that $R_1 = aR'$ and $R_2 = R'/a$. Then

$$\frac{V_{out}}{V_{in}} = \frac{1 + j\omega CR'/a}{1 + j\omega CaR'} \quad \text{from which} \quad \left| \frac{V_{out}}{V_{in}} \right| = \left\{ \frac{\left(\frac{1}{\omega CR'} \right)^2 + \frac{1}{a^2}}{\left(\frac{1}{\omega CR'} \right)^2 + a^2} \right\}^{1/2}$$

Rationalizing,

$$\begin{aligned} \frac{V_{out}}{V_{in}} &= \frac{(1 + j\omega CR'/a)(1 - j\omega CaR')}{1 + \omega^2 C^2 a^2 R'^2} \\ &= \frac{(1 + \omega^2 C^2 R'^2) + j(\omega CR'/a - \omega CaR')}{1 + \omega^2 C^2 a^2 R'^2} \\ \therefore \phi &= \tan^{-1} - \frac{\omega CR'a - \omega CR'/a}{1 + \omega^2 C^2 R'^2} \\ &= \tan^{-1} - \frac{a - \frac{1}{a}}{\frac{1}{\omega CR'} + \omega CR'} \end{aligned}$$

High-pass type—This is shown in *Figure 3.43*. Again, it differs from the

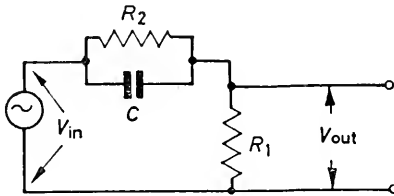


Figure 3.43

simple high-pass filter only in the addition of an extra resistance. By inspection,

$$\frac{V_{out}}{V_{in}} = \frac{R_1}{R_1 + \frac{R_2 \cdot \frac{1}{j\omega C}}{R_2 + \frac{1}{j\omega C}}}$$

which simplifies to

$$\frac{j\omega CR_1R_2 + R_1}{j\omega CR_1R_2 + R_1 + R_2}$$

Now assume, as is usually possible, that $R_2 \gg R_1$

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{j\omega CR_1R_2 + R_1}{j\omega CR_1R_2 + R_2}$$

Now let $R_2 = aR'$ and $R_1 = R'/a$

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{j\omega CR'^2 + R'/a}{j\omega CR'^2 + aR'} = \frac{j\omega CR' + \frac{1}{a}}{j\omega CR' + a}$$

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \left\{ \frac{(\omega CR')^2 + \frac{1}{a^2}}{(\omega CR')^2 + a^2} \right\}^{1/2}$$

Rationalizing, we find the phase shift is

$$\phi = \tan^{-1} \frac{a - \frac{1}{a}}{\frac{1}{\omega CR'} + \omega CR'}$$

The performance of these 'limited phase shift' filters is plotted in *Graphs 12* and *13* for various values of a . Evidently the price one has to pay for limited phase shift is limited attenuation.

R-C circuits which phase shift but do not filter

Phase-shifting networks are occasionally useful; they may be employed in conjunction with other circuitry to produce time delays, and are also

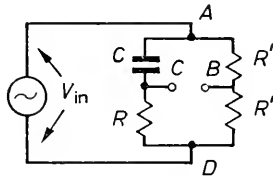


Figure 3.44

valuable in circuits for producing a 50 cycle signal of variable amplitude and phase for cancelling interference from mains pickup in amplifiers.

One such is shown in *Figure 3.44*; we have

$$\frac{V_{C-D}}{V_{\text{in}}} = \frac{j\omega CR}{1 + j\omega CR}$$

$$\frac{V_{B-D}}{V_{\text{in}}} = \frac{1}{2}$$

$$\frac{V_{B-C}}{V_{\text{in}}} = \frac{1}{2} - \frac{j\omega CR}{1 + j\omega CR}$$

RESISTANCES AND CAPACITANCES

$$= \frac{(1 + j\omega CR) - 2j\omega CR}{2(1 + j\omega CR)} = \frac{1(1 - j\omega CR)}{2(1 + j\omega CR)}$$

$$\left| \frac{V_{BC}}{V_{in}} \right| = \frac{1}{2} \frac{\{1 + (\omega CR)^2\}^{1/2}}{\{1 + (\omega CR)^2\}} = \frac{1}{2}$$

i.e. constant and independent of frequency. Rationalizing

$$\frac{V_{B-C}}{V_{in}} = \frac{(1 - j\omega CR)(1 - j\omega CR)}{1 + (\omega CR)^2} = \frac{1 - (\omega CR)^2 - 2j\omega CR}{1 + (\omega CR)^2}$$

$$\therefore \phi = \frac{-2\omega CR}{1 + (\omega CR)^2}$$

Remembering the trigonometrical half-angle formulae,

$$\tan \frac{\phi}{2} = -\omega CR$$

Thus ϕ can have any value between 0 and 180 degrees, for any given ω , as CR goes from 0 to ∞ .

A more generally useful version of this circuit is shown in *Figure 3.45*. It

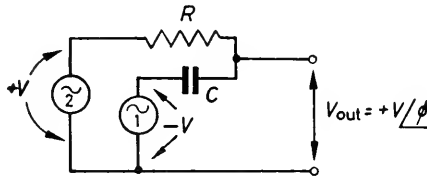


Figure 3.45

requires two inputs, equal and in anti-phase, such as may be obtained from a ‘concertina’ phase-splitter valve, to be described later in this part. The modulus of the output is V at all frequencies, and the phase shift is $\tan^{-1} \omega CR$, referred to the phase of generator 2.

We now come to consider a group of three, rather more elaborate, R–C networks; they are filters. One will pass all frequencies within a ‘band’, and attenuate those outside it—a ‘band-pass’ filter. The other two pass all frequencies *except* a band, which they attenuate; within the attenuation band there is a centre frequency at which the attenuation is infinite, and the networks are hence known as ‘null-transmission networks’. These are the Wien bridge and the parallel T network.

The R–C band-pass filter

This is shown in *Figure 3.46*. When presented with a new and complicated network it is a good plan not to plunge at once into calculations, but rather to get some idea what the network is likely to do. This is achieved by considering what happens when ω is very low, when ω is very high, and when ω is intermediate.

RESISTANCE AND CAPACITANCE IN SERIES

Thus, in the present case, when ω is very high the reactance of C_2 will become small compared with R_1 and may be neglected in comparison with it. Moreover, the reactance of C_1 will be small compared with R_2 , but as these elements are in parallel the one with the lower impedance is the more important. Hence, R_2 may be neglected in comparison with the effect of C_1 .

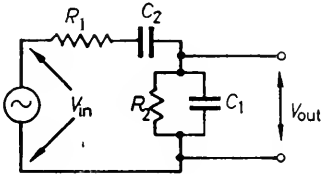


Figure 3.46

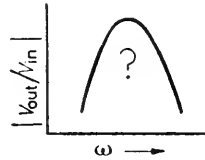


Figure 3.47

At ω large, then, the circuit simplifies to R_1 and C_1 , a simple low-pass filter, and V_{out} will fall with increasing frequency. By a similar process, we can deduce that when ω is low, the important elements are C_2 and R_2 , forming a simple high-pass filter, with V_{out} falling as the frequency is reduced. If the output is falling at both ends of the frequency spectrum, it presumably has a maximum at intermediate frequencies and the transmission characteristic has probably some such shape as Figure 3.47.

This, in fact, proves to be the case. Re-drawing the network in terms of

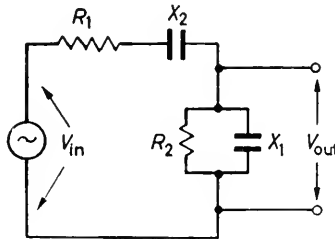


Figure 3.48

reactances instead of capacitances gives Figure 3.48. To prevent the investigation taking too long we shall set $X_1 = X_2 = X$, and see what happens as we vary R_1 and R_2 .

By inspection we can write down

$$\frac{V_{out}}{V_{in}} = \frac{\frac{-jR_2X}{R_2 - jX}}{R_1 - jX - \frac{jR_2X}{R_2 - jX}}$$

which simplifies to

$$\frac{V_{out}}{V_{in}} = \frac{-jR_2X}{(R_1R_2 - X^2) - jX(2R_2 + R_1)}$$

Now choose R , the geometric mean of R_1 and R_2 , and define a such that $R_2 = aR$ and $R_1 = R/a$. Then

$$\begin{aligned} \left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| &= \frac{-jaRX}{(R^2 - X^2) - jXR \left(2a + \frac{1}{a} \right)} \\ &= \frac{-ja}{\left(\frac{R}{X} - \frac{X}{R} \right) - j \left(2a + \frac{1}{a} \right)} \end{aligned}$$

Putting X equal to $1/\omega C$, we get the modulus of the transmission factor

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{a}{\left\{ \left(\omega CR - \frac{1}{\omega CR} \right)^2 + \left(2a + \frac{1}{a} \right)^2 \right\}^{1/2}}$$

This function is plotted in *Graph 14*, which may be used to design a filter for a given band-width centred on a given frequency. At the centre frequency the transmission factor rises from $1/3$, at $a = 1$, asymptotically towards $1/2$ as a approaches infinity.

Rationalizing to find the phase shift produced by the filter

$$\begin{aligned} \frac{V_{\text{out}}}{V_{\text{in}}} &= \frac{-ja \left[\left(\frac{R}{X} - \frac{X}{R} \right) + j \left(2a + \frac{1}{a} \right) \right]}{\left[\left(\frac{R}{X} - \frac{X}{R} \right) - j \left(2a + \frac{1}{a} \right) \right] \left[\left(\frac{R}{X} - \frac{X}{R} \right) + j \left(2a + \frac{1}{a} \right) \right]} \\ &= \frac{\left(2a + \frac{1}{a} \right) - ja \left(\frac{R}{X} - \frac{X}{R} \right)}{\left(\frac{R}{X} - \frac{X}{R} \right)^2 + \left(2a + \frac{1}{a} \right)^2} \\ \therefore \phi &= \tan^{-1} \frac{-a \left(\frac{R}{X} - \frac{X}{R} \right)}{2a + \frac{1}{a}} = \tan^{-1} \left[\frac{\frac{1}{\omega CR} - \omega CR}{2 + \frac{1}{a^2}} \right] \end{aligned}$$

(*Graph 15*)

Wien bridge

At $\omega CR = 1$, the R-C band-pass filter delivers a maximum output of modulus $V_{\text{out}} = \{a/(2a + 1/a)\}V_{\text{in}}$ with zero phase shift. If we connect across the input terminals a potential divider whose output is also $\{a/(2a + 1/a)\}V_{\text{in}}$, then the potential difference between the terminals A-C is zero at $\omega CR = 1$, and finite at other frequencies (*Figure 3.49*). The device is called a Wien bridge, and appears more bridge-like if re-drawn as in *Figure 3.50*. The Wien bridge is a null-transmission network.

RESISTANCE AND CAPACITANCE IN SERIES

In Figure 3.50 we have

$$\begin{aligned} \frac{V_{out}}{V_{in}} &= V_{AB} - V_{CB} \\ &= \frac{-ja}{\left(\omega CR - \frac{1}{\omega CR}\right) - j\left(2a + \frac{1}{a}\right)} - \frac{a}{2a + \frac{1}{a}} \end{aligned}$$

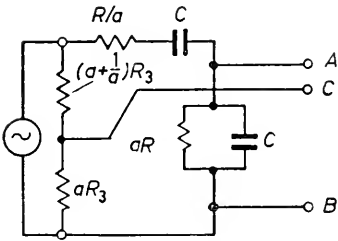


Figure 3.49

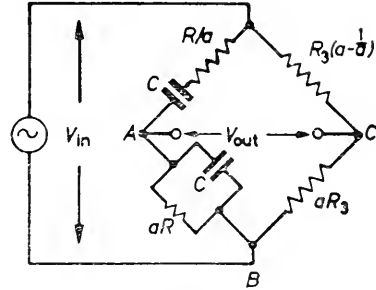


Figure 3.50

which simplifies to

$$\frac{V_{out}}{V_{in}} = \frac{a}{2a + \frac{1}{a}} \left\{ \frac{1}{j \frac{2a + 1/a}{\omega CR - \frac{1}{\omega CR}} - 1} \right\}$$

so

$$\left| \frac{V_{out}}{V_{in}} \right| = \frac{a}{2a + \frac{1}{a}} \left[\frac{1}{\left\{ \left(\frac{2a + \frac{1}{a}}{\omega CR - \frac{1}{\omega CR}} \right)^2 + 1 \right\}^{1/2}} \right]$$

Rationalizing to find the phase shift

$$\frac{V_{out}}{V_{in}} = \frac{1 + j \frac{2a + \frac{1}{a}}{\omega CR - \frac{1}{\omega CR}}}{1 + \left(\frac{2a + \frac{1}{a}}{\omega CR - \frac{1}{\omega CR}} \right)^2}$$

so

$$\phi = \tan^{-1} \frac{2a + \frac{1}{a}}{\omega CR - \frac{1}{\omega CR}}$$

RESISTANCES AND CAPACITANCES

Wien bridge transmission characteristic and phase shift are plotted for $a = 1$ and $a = 8$ in *Graphs 16* and *17*. Evidently $a = 1$ gives the sharper null, and in practical bridges this value is usually employed. In this case the transmission characteristic reduces to

$$\left| \frac{V_{out}}{V_{in}} \right| = \frac{1}{3} \frac{1}{\left(\left(\frac{3}{\omega CR - \frac{1}{\omega CR}} \right)^2 + 1 \right)^{1/2}}$$

and the phase shift to

$$\phi = \tan^{-1} \frac{3}{\omega CR - \frac{1}{\omega CR}}$$

Parallel T filter

It will become clear when we come to discuss valve circuits that a network of the form of *Figure 3.49* is not as convenient as one such as *Figure 3.51*, where one terminal of the output is directly connected to one terminal of the input, and both can, in practice, be earthed. *Figure 3.51* is called the parallel *T* filter, and has a performance similar to, but is rather more useful than, the Wien bridge.

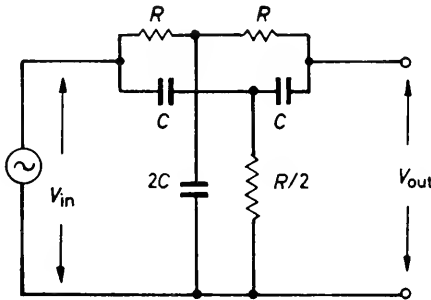


Figure 3.51

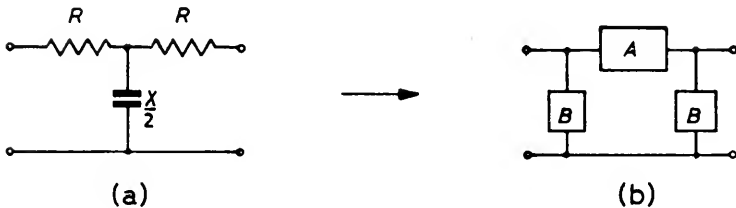


Figure 3.52

A good way to analyse this circuit is to convert each *T* into an equivalent pi; taking first the *T* possessing shunt capacitance (*Figure 3.52a*) and converting it to the pi in *Figure 3.52b* we get

$$A = R + R + \frac{R \cdot R}{-j \frac{X}{2}} = \frac{2R}{X} (X + jR)$$

RESISTANCE AND CAPACITANCE IN SERIES

$$B = R + \frac{-jX}{2} + R \cdot \frac{-jX/2}{R} = R - jX$$

Similarly, for the other T (Figure 3.53a), we get for the pi form (Figure 3.53b)

$$C = -jX - jX + \frac{-X^2}{R/2} = \frac{2X}{R} (X + jR)$$

$$D = -jX + \frac{R}{2} + \frac{-jX \cdot R/2}{-jX} = R - jX$$

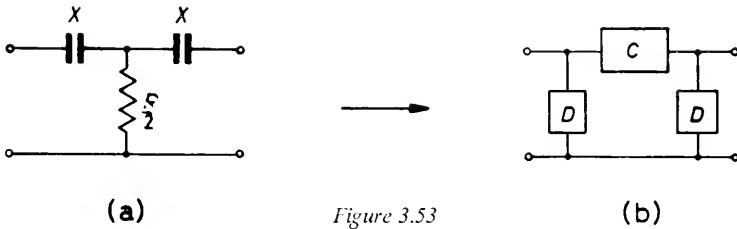


Figure 3.53

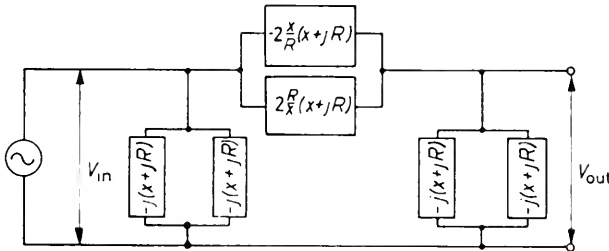


Figure 3.54

We now notice that B and D are both $R - jX$, and that if we take out $-j$ as a factor we get B and D both equal to $-j(X + jR)$. We now fit the two pi's together again (Figure 3.54) and by inspection

$$\begin{aligned} \frac{V_{out}}{V_{in}} &= \frac{B \text{ in parallel with } D}{B \text{ in parallel with } D + A \text{ in parallel with } C} \\ &= \frac{-\frac{1}{2}j}{2 \frac{R}{X} - 2 \frac{X}{R}} \\ &= \frac{-\frac{1}{2}j + \frac{2 \frac{R}{X} - 2 \frac{X}{R}}{2 \frac{R}{X} - 2 \frac{X}{R}}}{2 \frac{R}{X} - 2 \frac{X}{R}} \\ &= \frac{1}{1 - \frac{4j}{\frac{R}{X} - \frac{X}{R}}} \quad \text{Putting } X = \frac{1}{\omega C} \end{aligned}$$

$$\left| \frac{V_{out}}{V_{in}} \right| = \frac{1}{\left(1 + \left(\frac{4}{\omega CR - \frac{1}{\omega CR}} \right)^2 \right)^{1/2}}$$

(Graph 18)

and by the usual process the phase shift turns out to be

$$\phi = \tan^{-1} \frac{4}{\omega CR - \frac{1}{\omega CR}} \quad (\text{Graph 19})$$

Relationship between transmission characteristic and phase shift

It often happens that apparatus contains filters, either of the R-C type we have been discussing or of the R-C-L type to be dealt with in Chapter 5, whose precise natures are not known. If the transmission characteristic is known, or can be measured, then it is possible* to compute the phase shift.

If on the transmission characteristic we are considering a region of the curve which is reasonably straight, then the phase shift ϕ at frequency ω is given simply by

$$\phi_\omega = \frac{\pi}{12} \left(\frac{dA}{du} \right)_\omega$$

where $(dA/du)_\omega$ is the slope of the characteristic in dB's per octave at frequency ω . Thus a tapered 3-section low-pass R-C filter, well above the turn-over frequency, has a slope of 18 dB's per octave and the phase shift is

$$\phi_\omega = \frac{\pi}{12} \times 18 \quad \text{or} \quad 270 \text{ degrees}$$

If the transfer characteristic is rather less simple, perhaps a low-pass type of curve but containing a 'bump', as shown in *Figure 3.55*, then in the region

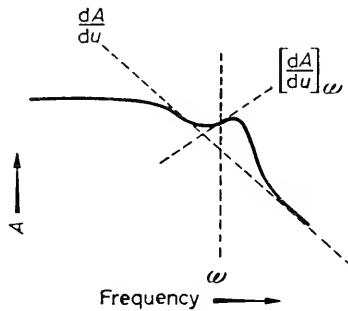


Figure 3.55

of the bump the phase shift is modified and it is necessary to add a correcting term to the expression already given. The equation which follows is due to Bode, and is quoted by F. E. Terman (*Radio Engineer's Handbook*). If the slope of the curve at the frequency ω where the phase shift is required is $(dA/du)_\omega$, and the general slope of the part of the characteristic upon which the bump is superposed is $(dA)/(du)$, as before, then

$$\phi_\omega = \frac{\pi}{12} \left(\frac{dA}{du} \right)_\omega + \frac{1}{6\pi} \int_{-\infty}^{+\infty} \left[\left(\frac{dA}{du} \right) - \left(\frac{dA}{du} \right)_\omega \right] \log_e \coth \left| \frac{u}{2} \right| du$$

where $u = \log_e (\text{frequency}/\text{frequency } \omega)$.

* Provided there are no 'all pass' sections present, e.g. 3.44 or 3.45.

RESISTANCE AND CAPACITANCE IN SERIES

A less fearsome expression which deals with the area under the phase-shift curve, rather than the phase shift itself, is

$$\int_{-\infty}^{+\infty} \phi_{\omega} \, d\omega = \frac{\pi}{17.37} (A_{\infty} - A_0)$$

where $u = \log_e \omega/\omega_r$, ω_r is any convenient reference frequency, and $(A_{\infty} - A_0)$ is the difference between the transmission factors of the filter at infinite and at zero frequency, measured in dB's. Thus in the filters whose transmission characteristic exhibits symmetry about a vertical axis—the Wien bridge and the parallel T — A_{∞} is equal to A_0 and the value of the integral is zero; hence the phase characteristic lies equally above and below the line $\phi = 0$. Again, in limited phase-shift filters, the equation shows that for a given degree of filtering, $A_{\infty} - A_0$, there is an inverse relationship between the phase shift encountered and the length of the transition band (i.e. the band over which the transmission factor is changing). In other words, one cannot have filters which cut-off sharply and exhibit small phase shift at the same time.

INDUCTANCES AND RESISTANCES

The magnetic circuit

When a current I flows through a coil of insulated wire, a magnetic field is set up which may be visualized as lines of force threading the coil, emerging at one end, passing outside the coil to the other end, and re-entering there. If the direction of current flow and the direction of the winding are as in *Figure 4.1*, the lines of force outside the coil pass up the paper as shown, and there is a north-seeking pole at the bottom of the coil, and a south-seeking pole at the top. The strength of the magnetic field is proportional to the current, to the number of turns, and to factors depending on the geometry of the coil. The total number of lines of force, or flux, is

$$\phi_{\text{air}} = \frac{K_1}{K_2} NI$$

where K_1/K_2 is the proportionality factor and N is the number of turns.

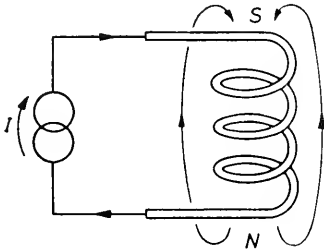


Figure 4.1

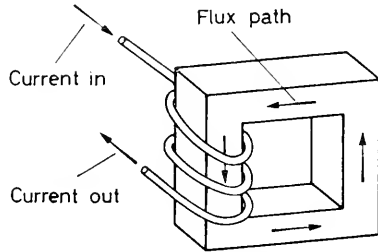


Figure 4.2

More lines of force are brought into being if they are provided with a path of low magnetic resistance or 'reluctance' to travel in. Thus if a closed ring or 'core' of iron or iron compound be provided, nearly all the flux is found to lie in the 'magnetic circuit' so produced, and to be proportional to a factor called μ , the 'permeability' of the iron (often of the order of 100 or 1,000), and inversely proportional to the length of the magnetic circuit (*Figure 4.2*)

$$\phi_{\text{iron}} = \frac{\mu K_1 NI}{l K_2}$$

Magnetic Ohm's law—If we compare the equation above with that for an elementary electrical circuit, $I = (1/R)E$, we see that if current is analogous to flux, then E , the electromotive force, is analogous to $K_1 NI$, the 'magnetomotive force', and $1/R$ is analogous to $\mu/l K_2$. Thus, $K_2 l / \mu$ corresponds to R , and is the *reluctance* of the magnetic circuit.

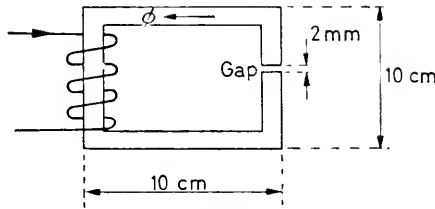
Air gap—It often happens that a magnetic circuit is provided with an

INDUCTANCES AND RESISTANCES

'air gap'. One effect of this is to make the flux for a given K_1NI rather independent of variations of μ which may occur in the iron, whilst retaining the valuable property of iron as a flux-guide.

The magnetic path-length in the arrangement of *Figure 4.3* is not far from 40 cm, whilst the length of the air gap is 2 mm. Let the μ of the iron be known to lie in the region 800–1,200: the permeability of air is unity: then we have

$$\phi = \frac{\text{m.m.f.}}{\text{Reluctance}}$$



If there were no gap, the reluctance would be proportional to the length of the flux path and inversely proportional to the iron permeability

$$\text{Reluctance}_{\text{no gap}} = K_2 \frac{l_{\text{iron}}}{\mu_{\text{iron}}} = K_2 \frac{40 \text{ cm}}{\mu_{\text{iron}}}$$

but with the air gap, by analogy with resistances in series

$$\begin{aligned} \text{Reluctance}_{\text{with gap}} &= K_2 \left(\frac{l_{\text{iron}}}{\mu_{\text{iron}}} + \frac{l_{\text{gap}}}{\mu_{\text{gap}}} \right) \\ &= K_2 \left(\frac{39.8 \text{ cm}}{\mu_{\text{iron}}} + \frac{0.2 \text{ cm}}{\mu_{\text{air}}} \right) \quad \text{but } \mu_{\text{air}} = 1 \\ &= K_2 \left(\frac{39.8}{\mu_{\text{iron}}} + 0.2 \right) \end{aligned}$$

If, for some reason, μ changes from 800 to 1,200, the change in reluctance without a gap would be

$$\left(\frac{K_2 \frac{40}{800}}{K_2 \frac{40}{1200}} \times 100\% \right) - 100\% = 50\%$$

but if μ changes from 800 to 1,200 with the air gap, the change in reluctance, and therefore of flux, is now only

$$\begin{aligned} &\left[\frac{K_2 \left(\frac{39.8}{800} + 0.2 \right)}{K_2 \left(\frac{39.8}{1200} + 0.2 \right)} \times 100\% \right] - 100\% \\ &\approx \left[\frac{0.05 + 0.2}{0.033 + 0.2} \times 100\% \right] - 100\% \approx 11\% \end{aligned}$$

Mutual inductance

When two coils of wire are arranged so that some of the flux caused by one threads the windings of the other, the arrangement is known as a 'mutual inductance'. It is found that an e.m.f. appears across the ends of one winding proportional to the rate of change of current in the other. In *Figure 4.4*, if

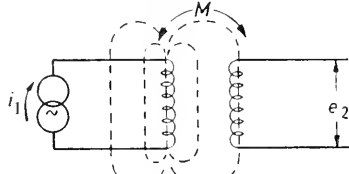


Figure 4.4

current changes in the primary winding at an instantaneous rate of one ampere per second, and if the e.m.f. induced into the secondary winding is one volt, then the value of the mutual inductance is one henry

$$e_2 = M di_1/dt \quad (\text{volts, henries, amps per second})$$

Mutual inductance is reciprocal, in the sense that, in addition,

$$e_1 = M di_2/dt$$

Self inductance

Since the flux produced by a coil threads the coil itself, it is not surprising that, upon varying the current, an e.m.f. appears across the terminals of the self-same coil. If the current in the winding is varied at the rate of one ampere per second, and if an e.m.f. of one volt appears across the coil terminals, then the 'self inductance' of the coil is one henry (*Figure 4.5*) and

$$e_1 = -L di/dt \quad (\text{volts, henries, amps per second})$$

The negative sign implies that the polarity of the induced e.m.f. is such as to oppose the flow of current if the latter is increasing, and vice versa; that is, it is a 'back e.m.f.'

The magnitude of the back e.m.f. is proportional to the rate of change of lines of force threading the circuit. By looping the circuit into a coil of N turns, the same flux is made to thread the circuit N times.

Thus $e = K_3 N d\phi/dt$

but $\phi = \mu \frac{K_1}{K_2} Ni$

so $\frac{d\phi}{dt} = \mu \frac{K_1}{K_2} N \frac{di}{dt}$

and $e = \frac{K_3 K_1 \mu}{K_2 l} N^2 \frac{di}{dt}$

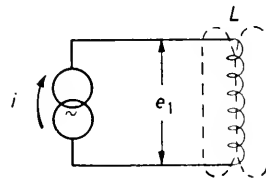


Figure 4.5

Comparing this with $e = -L di/dt$, it appears that $L = \{-(\mu K_3 K_1)/(K_2 l)\} N^2$ that is, the self inductance of a coil upon a given core is proportional to the

square of the number of turns. In Part II of the book we shall return to $(\mu K_3 K_1)/(K_2 l)$.

Self inductance in series

If two self inductances are connected in series to a generator whose current changes at a steady rate dI/dt (Figure 4.6) the back e.m.f. e_1 across L_1 is

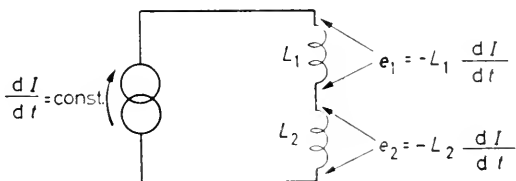


Figure 4.6

$-L_1 dI/dt$ and across L_2 is $e_2 = -L_2 di/dt$. The total back e.m.f. is therefore $-(L_1 + L_2) dI/dt$.

The back e.m.f. across a single coil of the same effective inductance L_{eff} would be $e = -L_{eff} dI/dt$

therefore
$$-L_{eff} dI/dt = -(L_1 + L_2) dI/dt$$

so
$$L_{eff} = L_1 + L_2$$

and in general for n self inductances in series $L_{eff} = L_1 + L_2 + L_3 \dots + L_n$.

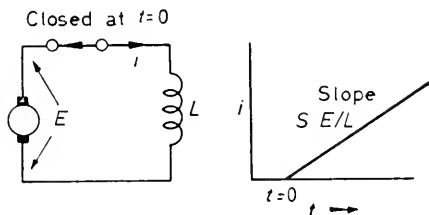


Figure 4.7

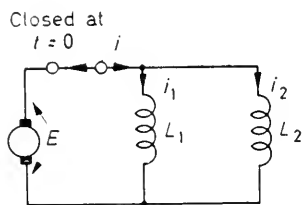


Figure 4.8

Self inductances in parallel

If a self inductance be connected to a constant direct voltage generator of e.m.f. E (Figure 4.7), the back e.m.f. e must necessarily be equal and opposite to the driving e.m.f. E , $E = -e$, and

$$E = L di/dt$$

therefore

$$di/dt = E/L$$

Thus from the instant of closing the switch the current rises from zero towards infinity at a steady rate. If, instead, two self inductances in parallel are connected to the generator (Figure 4.8) then

$$di_1/dt = E/L_1$$

and
$$di_2/dt = E/L_2$$

so
$$di/dt = di_1/dt + di_2/dt = E(1/L_1 + 1/L_2)$$

INDUCTANCES AND RESISTANCES

The rate of rise of current in a single coil of the same self inductance L_{eff} would be E/L_{eff} .

Therefore
$$E/E_{\text{eff}} = E(1/L_1 + 1/L_2)$$

and
$$1/L_{\text{eff}} = 1/L_1 + 1/L_2$$

and in general, for n self inductances in parallel,

$$\frac{1}{L_{\text{eff}}} = \frac{1}{L_1} + \frac{1}{L_2} + \frac{1}{L_3} \dots \frac{1}{L_n}$$

A more convenient form for two self inductances in parallel is

$$L_{\text{eff}} = \frac{L_1 L_2}{L_1 + L_2}$$

Self inductance connected to constant-voltage alternator

If a self inductance be connected to a generator of output $e = E \cos \omega t$ (Figure 4.9), there must be an equal and opposite back e.m.f. equal to $-E \cos \omega t$ produced across the inductance. Since this back e.m.f. is also $-L di/dt$, we have

$$L di/dt = E \cos \omega t$$

therefore

$$\begin{aligned} i &= (E/L) \int \cos \omega t dt \\ &= (E/\omega L) \sin \omega t \end{aligned}$$

Thus the current is also sinusoidal in form, but lagging in phase by 90 degrees on the applied voltage (Figure 4.10).

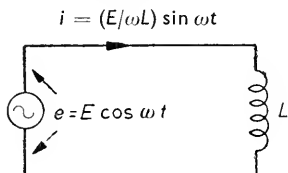


Figure 4.9

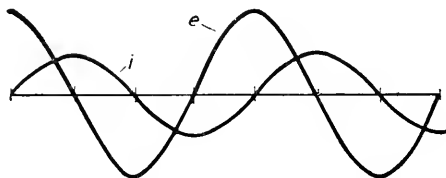


Figure 4.10

The ratio of applied voltage to current, neglecting the phase difference, is the modulus of the *inductive reactance*, X_L , and is $E/(E/\omega L) = \omega L$. Inductive reactance is conventionally regarded as positive. Thus, in j notation, taking account of the phase difference, the reactance is given by $j\omega L$.

The power delivered to the inductance per cycle is

$$\int_0^{2\pi} e \cdot i \cdot dt = \frac{E^2}{\omega L} \int_0^{2\pi} \sin \omega t \cdot \cos \omega t dt = 0$$

Thus an inductance, like a capacitance, consumes no power. Energy surges back and forth between the generator and the magnetic field of the inductance, but none is absorbed.

Self inductance in series with resistance, connected to constant direct voltage generator

When an inductance and a resistance are connected in series to a generator

INDUCTANCES AND RESISTANCES

of output E (Figure 4.11), the response is of the exponential type, for we have

$$E = V_R + V_L$$

If the instantaneous current is i , then $V_R = iR$, and $V_L = -e$, that is,

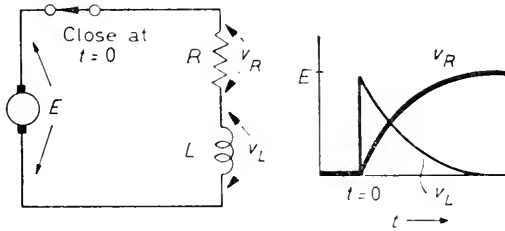


Figure 4.11

minus the back e.m.f. across the inductance, which is $+L \, di/dt$.

$$E = iR + L \, di/dt$$

for which the solution is

$$i = \frac{E}{R} (1 - e^{-(R/L)t})$$

Thus the current rises from 0 at $t = 0$, to a final value E/R with a time constant L/R . Since L/R is the time i would take to reach E/R if the initial rate of current rise were maintained, it follows that the initial rate of current rise is E/L , which is the same as the case for which the resistance is not present. The effect of R is to prevent the current rising steadily towards infinity, limiting it to the value E/R .

Self inductance and resistance in series, connected to a constant-voltage alternator

If R and L are connected to a generator of vector output V (Figure 4.12), the current which flows round the circuit is $V/(R + jX_L)$ and the potential

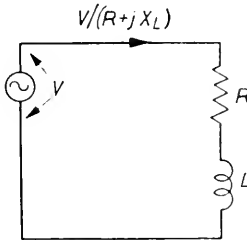


Figure 4.12

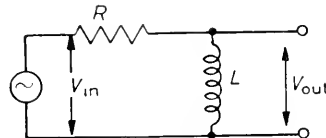


Figure 4.13

difference it produces across R is $(VR)/(R + jX_L)$ and across C is $(jVX_L)/(R + jX_L)$. These expressions could be the basis for another whole series of filters, for in the arrangement of Figure 4.13 we have

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{jX_L}{R + jX_L} = \frac{1}{1 + \frac{R}{jX_L}} \quad \left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{\left\{ \left(\frac{R}{\omega L} \right)^2 + 1 \right\}^{1/2}}$$

and for the arrangement of *Figure 4.14*

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{R}{R + jX_L} \quad \left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{\left\{ \left(\frac{\omega L}{R} \right)^2 + 1 \right\}^{1/2}}$$

Clearly these are of the same form as the equations $1/\{(1/\omega CR)^2 + 1\}^{1/2}$, for the simple R-C high-pass filter, and $1/\{(\omega CR)^2 + 1\}^{1/2}$ for the corresponding low-pass filter. That is, *Figure 4.13* is a high-pass filter which turns over at $\omega = R/L$, and *Figure 4.14* is low-pass, also turning over at $\omega = R/L$. R-L filters are not used in electronics, probably because inductors are more expensive, less available, bulkier, and generally less satisfactory than capacitors; but it is important to bear in mind that they exist, because they set a limitation to transformer performance, as we shall see.

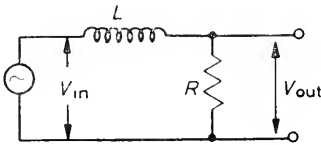


Figure 4.14

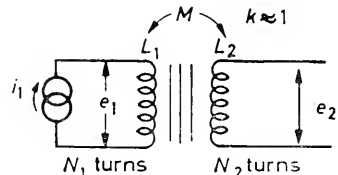


Figure 4.15

Tightly coupled mutual inductance

If an iron core be used to guide magnetic flux so that nearly all the lines of force due to current in one coil of a mutual inductance thread the turns comprising the other, then the two windings are said to be ‘tightly coupled’; this is in contradistinction to the state of affairs where the coils are some way apart, no core is provided, and the windings are said to be ‘loose coupled’. Most mutual inductances occurring in electrobiology are tightly coupled (i.e. transformers), but loose coupled circuits are employed when using radio frequency techniques (say, $\omega > 10^7$) as in the R-F coupled stimulator.

Suppose we have two tightly coupled windings, of self inductances L_1 and L_2 and mutual inductance M . If winding 1 be supplied with any varying current i_1 (*Figure 4.15*) then

$$e_1 = L_1 \frac{di_1}{dt}$$

and

$$e_2 = M \frac{di_1}{dt}$$

so

$$\frac{e_1}{e_2} = \frac{L_1}{M}$$

Now let the generator of i_1 be disconnected, and a generator of any other varying current i_2 be connected to winding 2, then

$$e_2 = L_2 \frac{di_2}{dt}$$

and

$$e_1 = M \frac{di_2}{dt}$$

so
$$\frac{e_1}{e_2} = \frac{M}{L_2}$$

Hence
$$\frac{e_1}{e_2} = \frac{L_1}{M} = \frac{M}{L_2}$$

\therefore
$$M = (L_1 L_2)^{1/2}$$

Turns ratio—If a tightly coupled mutual inductance be fed from some kind of generator as shown in *Figure 4.15*, we have

$$e_1 = K_3 N_1 d\phi/dt$$

and
$$e_2 = K_3 N_2 d\phi/dt$$

because of the tight coupling $\phi_1 = \phi_2$

so
$$e_1/e_2 = N_1/N_2$$

Loose coupled windings—the coupling factor k

If the coils of a mutual inductance are loosely coupled, M is less than $(L_1 L_2)^{1/2}$ and we introduce the factor k , such that $M = k(L_1 L_2)^{1/2}$. k then varies from zero—when the windings are indefinitely far apart, or are perpendicularly oriented—to nearly unity, when the coils are linked by an iron core and are close together.

Self inductances in series, but possessing some mutual coupling

If two coils, which considered separately have self inductances L_1 and L_2 are connected in series, their effective inductance is no longer merely $L_1 + L_2$ if there is mutual coupling between them (as when they share the same core); for let them be connected to a generator of any varying current i ; there is a back e.m.f. $-L_1 di/dt$ across L_1 due to the current in L_1 , and a back e.m.f. $-M di/dt$ across L_1 due to the current in L_2 . Similarly, there is a back e.m.f.

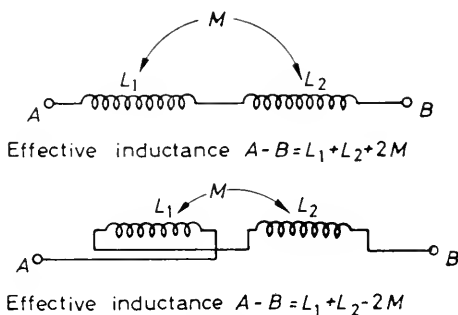


Figure 4.16

$-L_2 di/dt$ across L_2 due to the current in L_2 , and a back e.m.f. $-M di/dt$ across L_2 due to the current in L_1 . Thus the total back e.m.f. is $(L_1 + L_2 + 2M) di/dt$, hence the effective self inductance for the two windings is $L_1 + L_2 + 2M$. This assumes that the sense of the windings is such that their m.m.f.'s are acting in the same direction in the core. If the m.m.f.'s oppose one another, the effective self inductance is $L_1 + L_2 - 2M$ (*Figure 4.16*).

Self inductances in parallel, but possessing some mutual coupling

If two coils, which considered separately have self inductances L_1 and L_2 , are connected in parallel, their effective inductance is not $(L_1L_2)/(L_1 + L_2)$ if there is mutual coupling between them. It is not difficult to show that the effective inductance is now

$$L_{\text{eff}} = \frac{M^2 - L_1L_2}{2M - (L_1 + L_2)}$$

Generally speaking the only case of coupled inductances in parallel which is likely to occur is when $L_1 = L_2$. Putting M equal to $k(L_1L_2)^{1/2}$, we get

$$L_{\text{eff}} = \frac{L_1L_2(k^2 - 1)}{2k(L_1L_2)^{1/2} - (L_1 + L_2)}$$

and if $L_1 = L_2$

$$L_{\text{eff}} = \frac{L(1 + k)}{2}$$

Thus if $k = 0$, the effective inductance of the parallel combination is $L/2$, as anticipated. If $k = 1$, the effective inductance is merely L —a result which

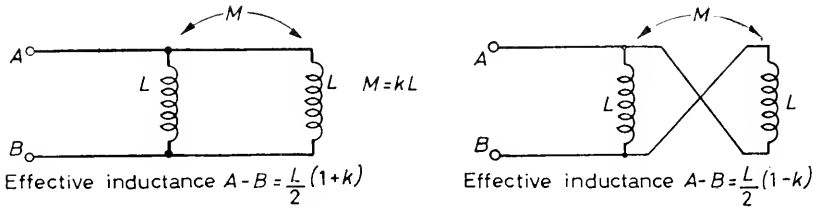


Figure 4.17

may at first sight seem surprising. This assumes that the sense of connections is such that the m.m.f.'s of the two coils act in the same direction. If they oppose one another the effective inductance for two equal coupled inductances becomes

$$L_{\text{eff}} = \frac{L(1 - k)}{2}$$

When $k = 0$, the effective inductance is $L/2$ as before. When $k = 1$, the effective inductance is zero (Figure 4.17)

Tightly coupled mutual inductance connected to a constant direct voltage generator

If a tightly coupled mutual inductance be connected to a generator of e.m.f. E (Figure 4.18), then the current rises in L_1 , as we have seen, according to the equation

$$di/dt = E/L_1$$

This current will induce in L_2 a constant e.m.f. $E_2 = M di/dt = (M/L_1)E_1$ but $M/L_1 = N_2/N_1$, therefore $E_2 = (N_2/N_1)E_1$.

INDUCTANCES AND RESISTANCES

This relationship is the basis for transformer action, for we have here an arrangement for transforming one e.m.f. into another which may be higher or lower than it.

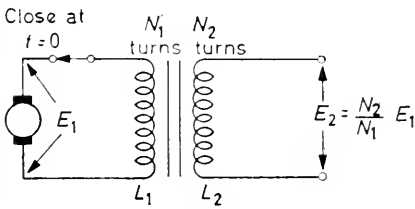


Figure 4.18

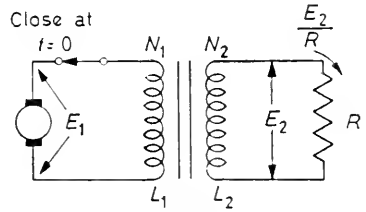


Figure 4.19

Tightly coupled mutual inductance connected between constant direct voltage generator and a load resistance

If now we connect a load resistance to L_2 in Figure 4.18, we get Figure 4.19. The current through the load resistance R will be E_2/R , and the power supplied to it E_2^2/R . This power has to come from the generator, and will be delivered by it in the form of extra current to L_1 . Thus we have

Extra generator power = Generator voltage \times Extra generator current

or
$$E_2^2/R = E_1 \times I_{G\text{-extra}}$$

therefore
$$I_{G\text{-extra}} = \frac{E_2^2}{E_1 R} = \frac{E_2}{E_1} \cdot \frac{E_2}{E_1} \cdot \frac{E_1}{R} = \left(\frac{N_2}{N_1}\right)^2 \frac{E_1}{R}$$

Thus, whereas with no load connected the generator supplies a current rising from 0 at the instant of switching on at the rate $dI/dt = E/L$ (Figure 4.20), if a load be connected then on closing the switch the generator current rises instantaneously by an amount $(N_2/N_1)^2 E_1/R$, and then continues to rise at the rate E/L (Figure 4.21).

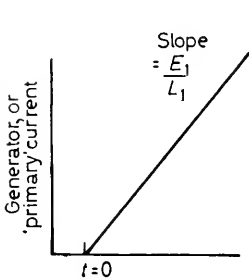


Figure 4.20

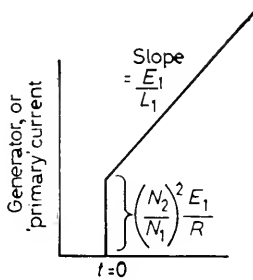


Figure 4.21

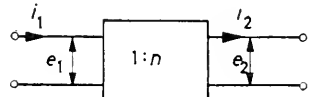


Figure 4.22

It is clear from the foregoing that the primary current comprises two fractions, a constant one—which represents energy being usefully transferred through to the load—and a steadily rising one—which represents energy being stored in the magnetic field, to no purpose. We therefore introduce the notion of the ideal transformer (Figure 4.22).

The ideal transformer

This is supposed to have a voltage ratio between input and output of n , so that $e_2 = ne_1$, and to be completely efficient, so that $e_2i_2 = e_1i_1$ and therefore $i_1 = ni_2$. If a load R be connected to the output terminals, then $i_2 = e_2/R$, and it is easy to see that $i_1 = n^2e_1/R$.

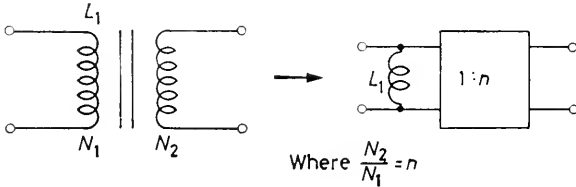


Figure 4.23

Comparing this with the transformation obtained by electromagnetic induction, we see that $i_1 = n^2e_1/R$ is analogous to the useful fraction of primary current $(N_2/N_1)^2 E_1/R$ where $1 : n :: N_1 : N_2$, and we can therefore represent the electromagnetic transformer by an ideal transformer and a self inductance (Figure 4.23).

Reflection through an ideal transformer— n^2E_1/R is the current which would flow if, instead of the ideal transformer and load resistance R , there was

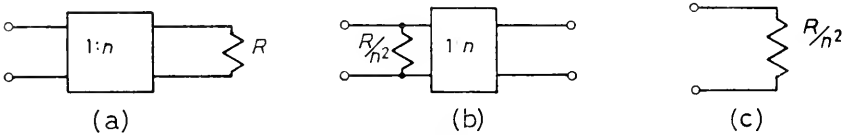


Figure 4.24

merely a resistance R/n^2 . It therefore appears that in Figure 4.24 a-c are equivalent. The secondary winding load resistance R is said to be reflected as a resistance R/n^2 across the input terminals.

Use of ideal transformer for matching—The ability of an ideal transformer to convert voltages and currents is often incidental to its ability to match a

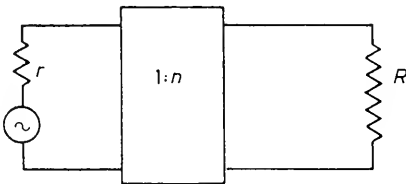


Figure 4.25

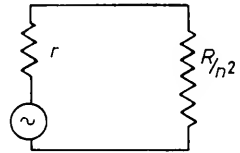


Figure 4.26

generator to a load. For in Figure 4.25, if we have a generator of internal resistance r , feeding a load resistance R , then R is reflected through the transformer as R/n^2 , and conditions for optimum power transfer are met when $n = (R/r)^{1/2}$ (Figure 4.26).

The real transformer

We are now in a position to carry out a simplified analysis of the behaviour of the real transformer, which is a practical component which attempts to emulate the ideal transformer in transforming voltages and currents, whether

direct or alternating, without loss of power. Real transformers are used in three ways.

Power transformer—Works on alternating current of a definite frequency and at a definite voltage from the mains, and steps the voltage up and down as required by other parts of the apparatus. Prime consideration, efficiency.

Pulse transformer—This handles a waveform such as *Figure 4.27*, and

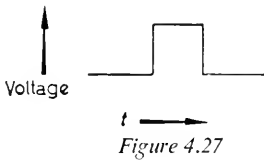


Figure 4.27

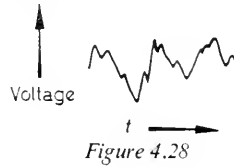


Figure 4.28

might be used, for example, to match the output of a stimulator to a preparation. Prime considerations, efficiency and absence of distortion.

Audio-frequency transformer—This deals with a 'signal', some complicated waveform such as *Figure 4.28*, which might be an E.C.G., E.E.G. or heart sound. A signal of this kind is composed of the sum of a number of pure sinusoidal waves of fluctuating frequency and amplitude, and the requirement of the transformer is that it should treat all the frequencies alike; thus prime consideration, absence of distortion.

At this point it may be objected that the pulse in *Figure 4.27* is also a signal, and that therefore a pulse transformer and a signal transformer are the same thing. In fact this is perfectly true, but the analyses for the two types are quite different because the questions one is likely to ask about their performances are different. We apply transient analysis to the pulse transformer, and steady-state analysis to the signal transformer.

In a real transformer we have two windings L_p and L_s linked by an iron core. The windings have resistances R_p and R_s because they are made of real wire and it cannot be avoided. The coupling is made as tight as it can possibly be made, yielding a k value usually better than 0.9. We represent

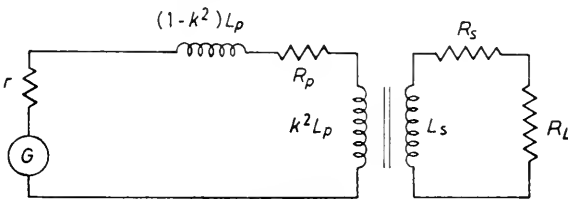


Figure 4.29

this state of affairs by splitting L_p into two parts, k^2L_p whose flux completely links with L_s , and $(1 - k^2)L_p$ whose flux fails to link with L_s . $(1 - k^2)L_p$ is called the 'leakage inductance' and k^2L_p the 'primary inductance'. For a good transformer k^2 is sufficiently near one for it to be permissible to call the primary inductance L_p . The arrangement feeds a load R_L and is fed by a generator of unspecified type having internal resistance r . This gives us *Figure 4.29*.

There ought to be a further resistance to represent losses occurring in the iron—hysteresis and eddy-current loss—but we cannot go too deeply into transformer theory here.

INDUCTANCES AND RESISTANCES

We proceed by simplifying *Figure 4.29*. Since inductance \propto turns² the turns ratio of the transformer is $(L_s/k^2L_p)^{1/2}$, but we shall not be far wrong if we call this $(L_s/L_p)^{1/2}$. We have seen that we can represent a mutual inductance by an ideal transformer and a self inductance; doing this in the case of *Figure 4.29* we get *Figure 4.30*.

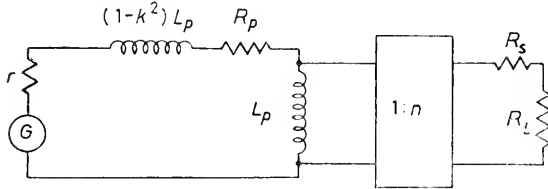


Figure 4.30

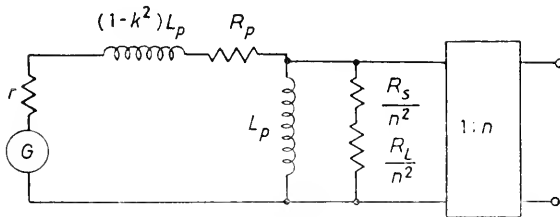


Figure 4.31

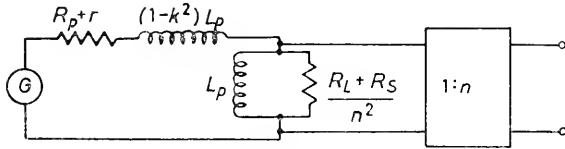


Figure 4.32

We now reflect R_s and R_L through the ideal transformer (*Figure 4.31*), and do a little more tidying to produce *Figure 4.32*. This is an equivalent circuit for a real transformer with a load. We now consider what happens when G takes specific forms.

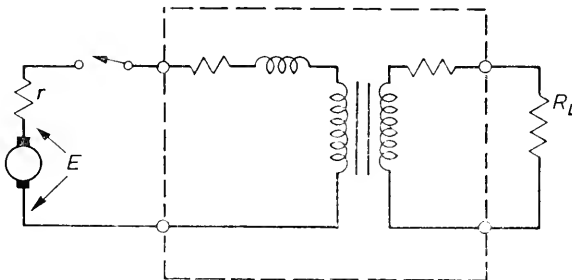


Figure 4.33

Loaded transformer connected to a constant-voltage type, direct voltage generator—the pulse transformer

If a real transformer is connected between a load R_L and a generator of direct e.m.f. E and internal resistance r (*Figure 4.33*), then to find out what

INDUCTANCES AND RESISTANCES

happens we need to know: (1) the turns ratio n ; (2) the coupling factor k ; (3) the primary inductance L_p ; and (4) the winding resistances. Armed with this information we construct the equivalent circuit (Figure 4.34). If the instantaneous currents after closing the switch are as shown

$$\text{also } \left. \begin{aligned} v_{AB} &= i_2 \left(\frac{R_s + R_L}{n^2} \right) \\ v_{AB} &= L_p \frac{d}{dt} (i_1 - i_2) \end{aligned} \right\} \frac{di_1}{dt} = i_2 \frac{R_s + R_L}{n^2 L_p} + \frac{di_2}{dt}$$

Further
$$E = i_1(R_p + r) + \frac{di_1}{dt} (1 - k^2) L_p + v_{AB}$$

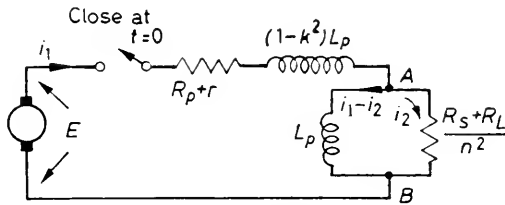


Figure 4.34

Combining the equations and differentiating gives

$$\frac{d^2 i_2}{dt^2} + \left\{ \frac{R_p + r + \frac{R_s + R_L}{n^2}}{L_p} + \frac{\frac{R_s + R_L}{n^2}}{(1 - k^2)L_p} \right\} \frac{di_2}{dt} + \frac{(R_p + r) \left(\frac{R_s + R_L}{n^2} \right)}{L_p^2 (1 - k^2)} = 0$$

The general solution for this equation is extremely cumbersome, and we shall not concern ourselves with it. There are three particular cases which are of interest, and these are:

- (1) $\frac{R_s + R_L}{n^2} \gg R_p + r$ which we might call the 'lightly loaded case';
- (2) $\frac{R_s + R_L}{n^2} = R_p + r$ which is approximately the matched condition;
- (3) $\frac{R_s + R_L}{n^2} \ll R_p + r$ the 'heavily loaded case', where the secondary winding is practically short-circuited.

The solutions are in the form of values of i_2 . More useful is v_{A-B} which is the input voltage to the ideal transformer and is $i_2(R_s + R_L)/n^2$. The actual voltage delivered to the load is then n times this.

Case (1) (lightly loaded)
$$v_{A-B} = E \cdot e^{-\frac{(R_p + r)t}{L_p}}$$

Case (2) (matched)
$$v_{A-B} = \frac{E}{2} \left(e^{-\frac{(R_s + R_L)t}{n^2 L_p}} - e^{-\frac{2(R_p + r)t}{(1 - k^2)L_p}} \right)$$

Case (3) (heavily loaded)

$$v_{A-B} = \frac{\frac{R_s + R_L}{n^2}}{(R_p + r) + \left(\frac{R_s + R_L}{n_2}\right)} \left(e^{-\frac{(R_s + R_L)t}{n^2 L_p}} - e^{-\frac{(R_p + r)t}{(1-k^2)L_p}} \right)$$

These are plotted in *Graph 20* for

Light load— $\frac{R_s + R_L}{n_2} = 10(R_p + r)$

Matched load— $\frac{R_s + R_L}{n_2} = R_p + r$

Heavy load— $\frac{R_s + R_L}{n_2} = 0.1(R_p + r)$

as a function of time for various coupling factors. Clearly a real transformer will not transmit a steady potential indefinitely, but it will do so for a short time, i.e. if the input is in the form of a pulse. If the loading is heavy the output pulse is small but is better sustained. The deficiencies of the output pulse may be described by ‘lag’, ‘sag’, and ‘undershoot’ (*Figure 4.35*); sag

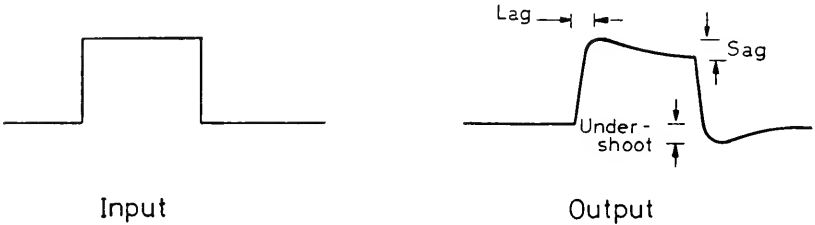


Figure 4.35

and undershoot are numerically equal. Summing up, for a transformer to transmit a pulse of a particular duration:

(1) If sag and undershoot are to be small, L_p must be large and, if a low output is acceptable, the load can be heavy.

(2) If the lag is to be small, k must be as near unity as possible, particularly if low sag has been achieved by heavy loading.

Loaded transformer connected to constant-voltage type alternator—signal and power transformer

The situation is depicted in *Figure 4.36*, but, as in the case of the pulse transformer, we work in terms of the equivalent circuit (*Figure 4.37*). Then by inspection,

$$\frac{V_{out}}{E_{in}} = \frac{\frac{R_L + R_s}{n_2} \left(jX_{L_p} / \frac{R_L + R_s}{n_2} + jX_{L_p} \right)}{r + R_p + j(1 - k^2)L_p + \frac{\frac{R_L + R_s}{n^2} jX_{L_p}}{\frac{R_L + R_s}{n^2} + jX_{L_p}}}$$

which, if k is nearly unity, simplifies to

$$\frac{V_{out}}{E_{in}} = \frac{1}{\left(\left(\frac{(1-k^2)\omega L_p}{R_L + R_s} - \frac{r + R_p}{\omega L_p} \right)^2 + \left(2 - k^2 + \frac{r + R_p}{R_L + R_s} \right)^2 \right)^{1/2}}$$

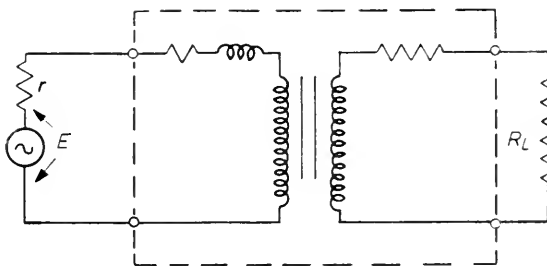


Figure 4.36

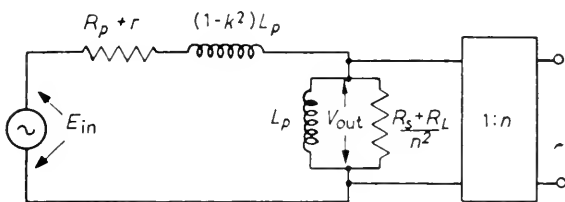


Figure 4.37

This is plotted as a function of frequency in *Graph 21* for various values of k and for 3 degrees of loading

Light load — $\frac{R_s + R_L}{n^2} = 10(R_p + r)$ —full lines

Matched load — $\frac{R_s + R_L}{n^2} = R_p + r$ —chain-dotted lines

Heavy load — $\frac{R_s + R_L}{n^2} = 0.1(R_p + r)$ —dashed lines

Evidently a transformer is a band-pass device having 6 dB/octave attenuation slopes at each end of the pass-band. The lower turn-over frequency occurs at

$$\omega_1 = \frac{(r + R_p) \left(\frac{R_L + R_s}{n^2} \right)}{L_p \left[(r + R_p) + \left(\frac{R_L + R_s}{n^2} \right) \right]}$$

so for a wide pass-band we need L_p large. The upper turn-over frequency is at

$$\omega_2 = \frac{(r + R_p) + \left(\frac{R_L + R_s}{n^2}\right)}{L_p(1 - k^2)}$$

so for a wide pass-band we need L_p small. Since the requirements for L_p conflict, we drop the notion of the actual turn-over frequencies and consider their ratio, since this gives us a measure of the band-width of the transformer.

$$\frac{\omega_2}{\omega_1} = \frac{\left[(r + R_p) + \left(\frac{R_L + R_s}{n^2}\right)\right]^2}{(1 - k^2)\left(\frac{R_L + R_s}{n^2}\right)(r + R_p)}$$

In the special but important case where the generator is matched to the load

$$\frac{R_s + R_L}{n^2} = R_p + r$$

and
$$\frac{\omega_2}{\omega_1} = \frac{4}{1 - k^2}$$

Summarizing for signal transformers under matching conditions:

- (1) The actual load voltage is nV_{out} (because our V_{out} has referred to the primary of the ideal transformer) and in the pass-band this equals $0.5 n E_{in}$, provided the winding resistances are small.
- (2) The bottom end of the pass-band is at $\omega = (r + R_p)/(2L_p)$
- (3) The pass-band extends over a $4/(1 - k^2)$ fold range of frequencies.

Danger—saturation

The equations which have been deduced to describe transformer performance have been linear, in the sense that it appears that in either pulse or signal transformer the input voltage can be indefinitely increased, to produce a corresponding increase in output voltage.

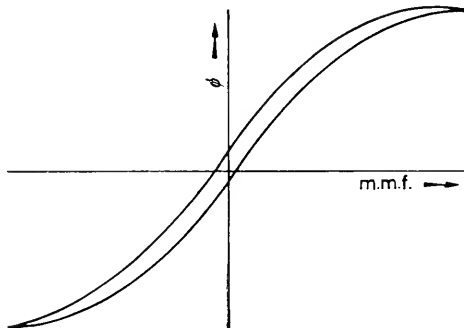


Figure 4.38

For transformers having iron cores this is not so. The linear theory assumes the primary inductance L_p to be a constant, and the inductance expression contains the permeability of the iron, μ , as a factor. If magneto-motive force is plotted against flux produced for a particular core, a curve results

INDUCTANCES AND RESISTANCES

which is closely related to the familiar B - H or hysteresis curve (*Figure 4.38*). The permeability is proportional to the slope of this, and is clearly not constant if the flux is allowed to exceed a certain limit in either direction.

Thus, for a given core, there is a definite allowable ϕ_{\max} if the linear analysis is to remain true. For any transformer

$$e_1 = K_3 N_1 \frac{d\phi}{dt}$$

For a pulse transformer, e_1 is a constant. Call it E . Then

$$E = K_3 N_1 \frac{d\phi}{dt}$$

and

$$\phi = \frac{Et}{K_3 N_1}$$

Therefore, if ϕ is not to exceed ϕ_{\max} , Et must not exceed $\phi_{\max} K_3 N_1$; that is, there is an upper limit to the amplitude-duration product of pulse which the transformer can transmit.

For a signal transformer, where the input is $V \cos \omega t$

$$\frac{d\phi}{dt} = \frac{V \cos \omega t}{K_3 N_1}$$

$$\therefore \phi = \frac{V}{\omega} \cdot \frac{\sin \omega t}{K_3 N_1}$$

If ϕ is not to exceed ϕ_{\max} , V/ω must not exceed $\phi_{\max} N_1 K_3$. Hence, for a given transformer fed with a given voltage, there is a lower frequency limit below which the output waveform becomes distorted due to core saturation. Matters can be improved again by reducing V . Since it is reasonable to fix the frequency at which distortion begins at about the lower turn-over frequency there will be associated with a particular transformer a maximum generator voltage with which it can be used if the full band-width is to be available.

In the case of power transformers, the voltages associated with the various windings are those which, if exceeded, will cause core saturation. Saturation in power transformers leads to very inferior efficiencies.

INDUCTANCES, CAPACITANCES AND RESISTANCES

Series L, C and R connected to a constant direct voltage generator

The behaviour of this relatively simple-looking circuit (*Figure 5.1*) is quite complicated, but well repays study; it is the electrical analogue of mechanical devices such as the electromagnetic penwriter, and an understanding of it is of great help in grasping how such devices perform.

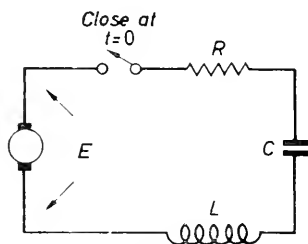


Figure 5.1

On closing the switch we have

$$\begin{aligned} E &= v_R + v_C + v_L \\ &= iR + \frac{1}{C} \int i \, dt + L \frac{di}{dt} \end{aligned}$$

but $i = dq/dt$, therefore

$$E = \frac{dq}{dt} R + \frac{q}{C} + L \frac{d^2q}{dt^2}$$

Let us consider the voltage across the capacitance, v_C . This is q/C , so

$$E = RC \frac{dv_C}{dt} + v_C + LC \frac{d^2v_C}{dt^2}$$

There are three types of solution to this equation, depending on the amount of resistance in the circuit. When the resistance is small, the circuit is 'lightly damped' and the response is oscillatory, the oscillations dying away slowly when R is very small, and more quickly as R is raised. This oscillatory response is called 'ringing', by analogy with the striking of a bell. On further increasing R we reach at length a point at which the oscillations die away, as it were, before they have started, and the response of the circuit is said to be 'critically damped'. If R be made greater still, the response is of the 'heavily damped' variety, extremely sluggish.

On solving the equation it turns out that everything depends on the quantities $1/LC$ and $R^2/4L^2$. When $R^2/4L^2 \ll 1/LC$, the solution is approximately

$$v_C = E(1 - e^{-Rt/2L}) \cos \frac{t}{(LC)^{1/2}}$$

To see how this looks, let $R^2/4L^2 = 0.1/LC$. Then the equation becomes

$$v_C = E(1 - e^{-t/(10LC)^{1/2}}) \cos \frac{t}{(LC)^{1/2}}$$

When $R^2/4L^2 = 1/LC$ the circuit is critically damped and the solution is

$$v_C = E \left[1 - e^{-t/(LC)^{1/2}} \left(1 + \frac{t}{(LC)^{1/2}} \right) \right]$$

When $R^2/4L^2 \gg 1/LC$, the circuit is heavily damped and the general solution is a somewhat cumbersome expression. However, in the case of $R^2/4L^2 = 10/LC$ the solution is

$$v_C = E(1 - 1.025 e^{-0.162t/(LC)^{1/2}} + 0.025 e^{-6.162t/(LC)^{1/2}})$$

These three responses are plotted in *Graph 22*, which shows the three routes by which the capacitance voltage eventually becomes equal to E . *Graph 22* also shows, though not rigorously, that *the critically damped circuit settles into the final state soonest*: this is a point of great importance.

A physical description of what is happening may not be out of place. In the lightly damped case, upon closing the switch, current tends to flow to charge the capacitance, but its growth is retarded by the back e.m.f. across the inductance, which opposes E . The initial slope of the current curve is zero. As charging proceeds a magnetic field is built up around the inductances. As v_C approaches E , the current would fall off were it not for the e.m.f. across the inductance, which now aids E , charging the capacitance to a voltage greater than E . Energy is being transferred from the magnetic field, which is collapsing, to the capacitance in the form of additional stored charge.

When the energy of the magnetic field is exhausted the capacitance voltage begins to push current round the circuit the other way, against the action of the generator, building up a magnetic field round the inductance once more and giving up its own energy to do so.

Thus, the oscillations represent the transfer of energy back and forth between the inductance and the capacitance. If R were not present this process could continue indefinitely. The effect of R is to absorb a fraction of the energy at each cycle, converting it irreversibly into heat, so that the oscillations die away. As R is increased they die away more and more quickly, until in the critically damped case they have been suppressed altogether. If R be made larger still, the circuit begins to turn into a simple R-C low-pass filter, because the effect of R swamps that of L . Hence, increasing R is increasing the time constant, and the response becomes increasingly sluggish.

Capacitive and inductive reactance

If series R, L, and C are connected to a constant alternating current generator of output $i = I \sin \omega t$ (*Figure 5.2*), then v_R is in phase with i , v_L leads

INDUCTANCES, CAPACITANCES AND RESISTANCES

on i by 90 degrees, and v_C lags on i by 90 degrees. It follows that v_L and v_C are in anti-phase with one another, that is, they pass through maxima and minima in step with each other, but with opposite polarity, as suggested in Figure 5.2.

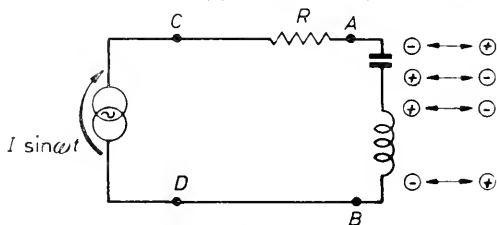


Figure 5.2

The instantaneous potential difference between A and B is v_L minus v_C , and the modulus of the reactance of the combination is $|v_L - v_C|/I$ which equals $|v_L|/I - |v_C|/I$. However $|v_L|/I = X_L$ and $|v_C|/I = X_C$ so the reactance is $X_L - X_C$. This is the reason for taking capacitive reactance as negative and inductive reactance as positive. Common sense suggests that if a number of reactances are connected in series, the effective reactance ought to be obtained by adding them all up. And so it can be, for

$$X_{\text{eff}} = X_{L1} + X_{L2} + X_{L3} \dots X_{Ln} + (-X_{C1}) + (-X_{C2}) \dots + (-X_{Cn})$$

The impedance of the arrangement between C and D is therefore

$$\begin{aligned} & R + j(X_L - X_C) \\ &= R + j\left(\omega L - \frac{1}{\omega C}\right) \end{aligned}$$

When ω is small, ωL is smaller than $1/\omega C$ and the circuit behaves as if it contained only the resistance and a capacitance C' such that $1/\omega C' = 1/\omega C - \omega L$. When ω is large, ωL is larger than $1/\omega C$ and the circuit behaves as if it contained only the resistance and the inductance L' such that $\omega L' = \omega L - 1/\omega C$. When ω has the critical value $1/(LC)^{1/2}$ the reactances cancel, and it is as if only the resistance was there. This is the basis of series resonance.

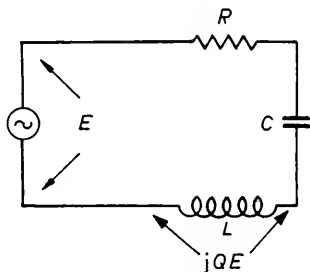


Figure 5.3

Series resonance

If series L, C, and R are connected to a constant voltage alternator (Figure 5.3), the current is given by

$$I = \frac{E}{R + j\left(\omega L - \frac{1}{\omega C}\right)}$$

and rises to a maximum E/R when $\omega L = 1/\omega C$, that is, at $\omega = 1/(LC)^{1/2}$. The current is then in phase with the applied voltage and the circuit is said to be 'series resonant'. At resonance the voltage across the inductance is given by

$$V_L = I \cdot j\omega L = \frac{E}{R} \cdot j\omega L$$

and
$$\frac{V_L}{E} = \frac{E}{R} \cdot \frac{j\omega L}{E} = \frac{j\omega L}{R} = jQ$$

$\omega L/R$ is the voltage magnification factor of the circuit and in practical cases may be over 100. It is allotted the symbol Q . Since at resonance $\omega = 1/(LC)^{1/2}$, Q is also given by

$$\frac{\{1/(LC)^{1/2}\}L}{R} = \left(\frac{L}{CR^2}\right)^{1/2}$$

Consider now the voltage across the capacitance. We have, by inspection

$$\frac{V_C}{E} = \frac{-j\omega C}{R + j\left(\omega L - \frac{1}{\omega C}\right)}$$

Since $Q = (L/CR^2)^{1/2}$ so that $Q/(L/CR^2)^{1/2} = 1$, we have that for the numerator

$$-j\omega C = -j\omega C \cdot \frac{Q}{(L/CR^2)^{1/2}} = \frac{-jQR}{\omega(LC)^{1/2}}$$

Since Q also equals $\omega_o L/R$, where ω_o is the resonant frequency, $QR/\omega_o L = 1$ and the denominator can be written

$$R \left\{ 1 + j \frac{QR}{\omega_o L} \left(\frac{\omega L}{R} - \frac{1}{\omega CR} \right) \right\}$$

therefore

$$\frac{V_C}{E} = \frac{-jQ}{\omega(LC)^{1/2} \left(1 + jQ \left[\frac{\omega}{\omega_o} - \frac{1}{\omega \cdot \omega_o LC} \right] \right)}$$

and letting $1/(LC)^{1/2} = \omega_o$

$$\frac{V_C}{E} = \frac{-jQ}{\frac{\omega}{\omega_o} \left\{ 1 + jQ \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) \right\}}$$

$$\left| \frac{V_C}{E} \right| = \frac{Q}{\frac{\omega}{\omega_o} \left\{ 1 + Q^2 \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 \right\}^{1/2}}$$

This is plotted against ω for various Q in *Graph 23*. Observe how $V_C = E$ at frequencies well below resonance, but that at the resonant frequency the circuit magnifies the generator voltage Q times. Above resonance V_C falls away rapidly.

The voltage across the inductance may be found, as a function of frequency, in a similar manner. The curves are of the same kind but are mirror images (Figure 5.4).

These curves have Q as parameter and $Q = (L/CR^2)^{1/2}$, that is, Q is large when R^2/L is small compared with $1/C$. Referring back to the analysis for series L , C and R connected to the direct voltage generator, we see that the

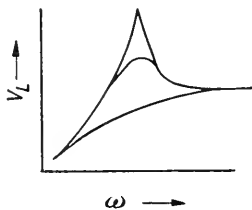


Figure 5.4

degree of damping depended on the relative sizes of $1/LC$ and $R^2/4L^2$, which is the same as depending on the sizes of $1/C$ and $R^2/4L$. It thus appears that Q and the degree of damping are closely related, and in particular that at critical damping $Q = \frac{1}{2}$; series resonant circuits are below critical damping if $Q > \frac{1}{2}$, and more than critically damped if $Q < \frac{1}{2}$.

We can now consider the relevance of *Graphs 21* and *22* to the electromagnetic penwriter. The input current flows through a coil to produce a force F on a moving mechanical system of mass M constrained by a spring of stiffness λ and viscous damping δ . If the displacement produced is s , then we have

$$F = \delta \frac{ds}{dt} + \lambda s + M \frac{d^2s}{dt^2}$$

and the actual force available for displacing the pen is $F_{\text{pen}} = \lambda s$, the remainder of F being used in overcoming the inertia and viscous damping forces. Comparing this with

$$E = R \frac{dq}{dt} + \frac{q}{C} + L \frac{d^2q}{dt^2}$$

where the capacitance voltage is $v_C = q/C$, we see that the pen deflection is analogous to capacitance voltage. Hence, if a current be suddenly applied to a penwriter, the pen movement has one of the forms of *Graph 21*, depending on the amount of viscous damping present. If an alternating current of constant amplitude and variable frequency be fed to a penwriter the device records, without distortion, up to a certain frequency, where there is a greater or lesser resonance, depending upon the amount of damping, and above this frequency the performance of the device falls off rapidly. From *Graph 22* it appears that a Q value of rather less than one will give the penwriter an optimum distortion-free band-width.

Impedance of series resonant circuit

The impedance of a series resonant circuit at frequencies away from resonance is important in the design of certain tuned amplifiers.

We have $Z_s = R + j\left(\omega L - \frac{1}{\omega C}\right)$

Let $R = \frac{1}{Q}\left(\frac{L}{C}\right)^{1/2}$

$$Z_s = \frac{1}{Q}\left(\frac{L}{C}\right)^{1/2} + j\left(\omega L - \frac{1}{\omega C}\right)$$

$$= \frac{1}{\omega C} \left\{ \frac{\omega}{Q} (LC)^{1/2} + j(\omega^2 LC - 1) \right\}$$

then, letting $\omega_o = \frac{1}{(LC)^{1/2}}$

$$Z_s = \frac{1}{\omega C} \left[\frac{\omega}{\omega_o} \cdot \frac{1}{Q} + j\left(\left(\frac{\omega}{\omega_o}\right)^2 - 1\right) \right]$$

$$= \frac{1}{\omega_o C} \left\{ \frac{1}{Q} + j\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right) \right\}$$

$$|Z_s| = \frac{1}{\omega_o C} \left\{ \frac{1}{Q^2} + \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)^2 \right\}^{1/2}$$

$1/\omega_o C$ is a constant for the circuit. In *Graph 24*, the impedance of a series resonant circuit is plotted in the form $|Z_s| \omega_o C$ as a function of frequency for various Q .

Parallel L-C-R circuit connected to constant-current generator

The parallel circuit of *Figure 5.5* is of importance because, if a capacitance be connected across a real inductor, then the real inductor is represented by an ideal inductance L in series with the resistance of the wire comprising the

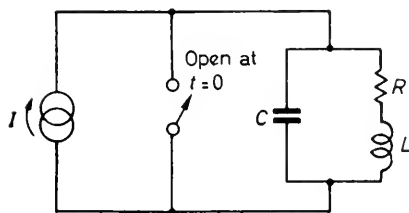


Figure 5.5

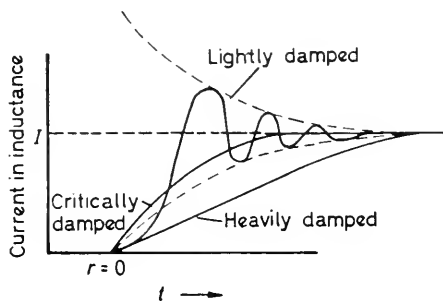


Figure 5.6

inductor, R . If this circuit be suddenly fed with a current I , it settles down to a final state in which all of I flows through L and R by one of three approaches—oscillatory, critically damped, or heavily damped (*Figure 5.6*) in a manner analogous to the capacitance voltage when series $L-C-R$ is connected to a constant-voltage generator. As in the series case, critical damping occurs at $R^2/4L^2 = 1/LC$.

Parallel resonance

Suppose the parallel L-C-R circuit be connected to a constant voltage alternator (*Figure 5.7*) and suppose $R = 0$. Then the impedance of the circuit is $\{j\omega L \cdot -j/\omega C\}/\{j\omega L - j/\omega C\}$. When $\omega = 1/(LC)^{1/2}$ the denominator becomes zero and the impedance goes to infinity (contrast this with the series circuit, in which the impedance at resonance is a minimum).

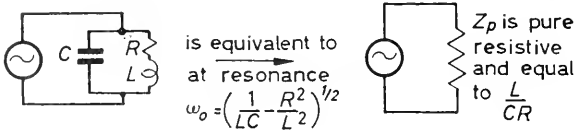


Figure 5.7

With a practical circuit $R \neq 0$ and at resonance the impedance does not become infinite. We have

$$Z_p = \frac{(R + j\omega L) \cdot -\frac{j}{\omega C}}{(R + j\omega L) - \frac{j}{\omega C}}$$

which simplifies and rationalizes to

$$Z_p = \frac{\frac{R}{C^2\omega^2} + j\left(\frac{L}{C^2\omega} - \frac{R^2}{C\omega} - \frac{\omega L^2}{C}\right)}{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}$$

Parallel resonance is defined as occurring at that frequency at which Z_p becomes purely resistive. Thus the j term in the above equation must disappear, i.e.

$$\frac{L}{C^2\omega_o} - \frac{R^2}{C\omega_o} - \frac{\omega_o L^2}{C} = 0$$

where ω_o is the parallel resonant frequency. The solution is

$$\omega_o = \left(\frac{1}{LC} - \frac{R^2}{L^2}\right)^{1/2}$$

Substituting back, it emerges eventually that, at ω_o , the impedance of the circuit is simply

$$Z_p = \frac{L}{CR}$$

L/CR is called the ‘dynamic resistance’ of the circuit.

Parallel L-C-R connected to constant-current alternator

In *Figure 5.8*, if the parallel circuit is resonant at the alternator frequency, then it may be replaced by a resistance L/CR and the potential difference across the circuit is $V = I \cdot L/(CR)$.

The capacitance current is $V/(-j/\omega C) = j\omega CV$.

INDUCTANCES, CAPACITANCES AND RESISTANCES

The inductance current is $V/(R + j\omega L)$ which, if R is fairly small (Q of circuit high), is approximately equal to $-jV/\omega L$. Also, if R is small, $R \ll (L/C)^{1/2}$ and the resonant frequency is approximately given by $\omega = 1/(LC)^{1/2}$ so that $\omega L \approx 1/\omega C$. The capacitance and inductance currents are

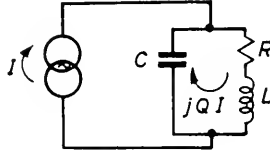


Figure 5.8

now seen to be equal and in opposite directions; a *circulating current* flows within the resonant circuit.

The circulating current is $j\omega CV = j\omega C \cdot I(L/CR)$

The *current magnification* of the circuit is

$$\begin{aligned} & \frac{\text{Circulating current}}{\text{Generator current}} \\ &= j\omega L/R \\ &= jQ \end{aligned}$$

Physically, the circulating current represents the transfer of energy back and forth between electric field in the capacitance and magnetic field around the inductance. At each transfer some energy is lost in the resistance, and this is replaced by a small current from the generator.

Impedance of parallel resonant circuit when off resonance

As for the series case, this function is important in the design of tuned amplifiers. We have

$$\begin{aligned} Z_p &= \frac{(j\omega L + R) \frac{1}{j\omega C}}{R + j\omega L + \frac{1}{j\omega C}} \\ &= \frac{j\omega L + R}{j\omega CR - \omega^2 LC + 1} \end{aligned}$$

Let

$$R = \frac{1}{Q} \left(\frac{L}{C} \right)^{1/2}$$

Then

$$\begin{aligned} Z_p &= \frac{j\omega L + \frac{1}{Q} \left(\frac{L}{C} \right)^{1/2}}{j\omega \frac{1}{Q} \left(\frac{L}{C} \right)^{1/2} - \omega^2 LC + 1} \\ &= \omega L \frac{jQ + \frac{1}{(LC)^{1/2}\omega}}{j\omega(LC)^{1/2} + Q(1 - \omega^2 LC)} \end{aligned}$$

Now let

$$\omega_o = \frac{1}{(LC)^{1/2}}$$

$$Z_p = \omega_o L \frac{jQ + \frac{\omega_o}{\omega}}{j + Q\left(\frac{\omega_o}{\omega} - \frac{\omega}{\omega_o}\right)}$$

At $\omega/\omega_o = 1$, this is approximately equal to $\omega_o LQ$ if Q is large. Putting ω_o back in terms of L and C , and Q back in terms of L , C , and R , we see that $\omega/\omega_o = 1$, $Z_p = L/(CR)$, as before.

Further, we have

$$|Z_p| = \omega_o L \left\{ \frac{\left(\frac{\omega_o}{\omega}\right)^2 + Q^2}{1 + Q^2\left(\frac{\omega_o}{\omega} - \frac{\omega}{\omega_o}\right)^2} \right\}^{1/2}$$

$\omega_o L$ is a constant of the circuit. The impedance function for a parallel resonant circuit is plotted in the form $|Z_p|/\omega_o L$ as a function of frequency, for various Q , in *Graph 25*.

Shunt damping

We often need to control the damping of a resonant circuit, and this is conveniently achieved by varying the amount of resistance present. Hitherto we have considered the damping resistance as being in series with the inductance, because this condition must obtain with real components, being the resistance of the inductor winding. However, if the damping on a given resonant circuit is insufficient we can get more by connecting additional resistance in series with the inductor, or by connecting resistance across it

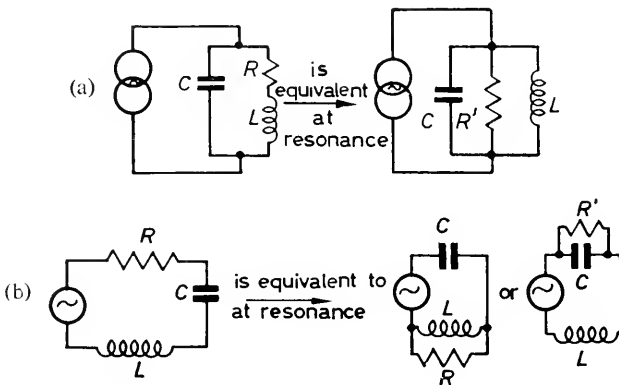


Figure 5.9

(or across the capacitance). When a resistance is connected *across* one of the reactive elements the circuit is said to be 'shunt damped'; thus the circuits of *Figure 5.9* are equivalent. The amount of shunt damping required is suggested by the fact that a parallel resonant, series-damped circuit containing resistance R behaves as a resistance $R' = L/CR$ (*Figure 5.9a*). Thus,

INDUCTANCES, CAPACITANCES AND RESISTANCES

if the circuit is critically series-damped by series resistance $R = 2(L/C)^{1/2}$, it will also be critically damped by a shunt resistance

$$R' = \frac{L}{C \left(2 \left(\frac{L}{C} \right)^{1/2} \right)} = \frac{1}{2} \left(\frac{L}{C} \right)^{1/2}$$

Conclusions

Generator	Series L-C-R	Parallel L-C-R
Constant-voltage alternator— E	At resonance, current rises to max. E/R . Voltage magnification of Q times available across reactive elements	At resonance, generator current falls to min. $E \propto (L/CR)$
Constant-current alternator— I	At resonance, terminal voltage falls to min. IZ_s according to <i>Graph 23</i>	At resonance, terminal voltage rises to maximum IZ_v according to <i>Graph 24</i> . Current magnification of Q times in circuit
Constant direct voltage	Varies from vigorous ringing to sluggish follow-up, depending on damping	Trivial
Constant direct current	Trivial	Varies from vigorous ringing to sluggish follow-up, depending on damping

Critical series damping, $R = 2(L/C)^{1/2}$. Critical shunt damping, $R = \frac{1}{2}(L/C)^{1/2}$.
Critical damping corresponds to $Q = 1/2$.

The reader may care to satisfy himself that the cases marked 'trivial' are in fact so. The circuit responses are the sum of responses found for more elementary circuits in previous chapters.

Effect of a series resonance upon the performance of signal transformers

In deriving the equivalent circuit for the signal transformer we have neglected the effects of the self-capacitance of the two windings.

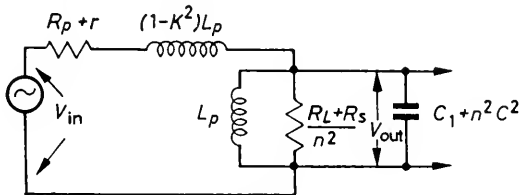


Figure 5.10

Let the effect of the distributed secondary capacitances be represented by a capacitor C_2 connected across the secondary winding, and let the distributed primary capacitances be represented by C_1 across the primary winding. Then C_2 is reflected across to the primary side as $n^2 C_2$ (Figure 5.10). At a

frequency which is usually in the region of the upper end of the pass-band a series resonance occurs between $C_1 + n^2C_2$ and $(1 - k^2)L_p$ which appears as a Q -fold increase in the input to the ideal transformer. The effect is to put a 'hump' into the transmission characteristic (*Figure 5.11*) whose

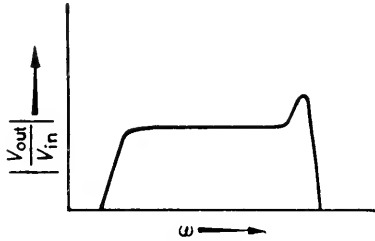


Figure 5.11

magnitude depends on the shunt damping $(R_L + R_s)/n^2$ and the series-damping $r + R_p$. With good design the hump can be useful in extending the pass-band at the upper frequency end.

This resonance sets a limit to the voltage step-up which may be had from a signal transformer; for if n is too great, $C_1 + n^2C_2$ becomes of such a size that the resonance comes down into the hoped-for pass-band. Since transmission falls away quickly above the resonance (*Graph 22*) the pass-band actually obtained will be seriously narrowed.

Smoothing section

In the language of resonant circuits, *Figure 5.12* is a shunt-damped, series-resonant circuit, and we may anticipate that it will have a transmission characteristic similar to the series-damped case (*Graph 23*). In the next

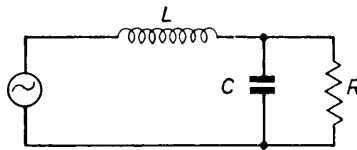


Figure 5.12

chapter we shall see that when power from the a.c. mains is rectified to produce a direct voltage, the output carries a superposed 'ripple'. This ripple is generally undesirable, and a 'smoothing section', comprising an inductor and a capacitor, is employed to filter it out. In *Figure 5.12*, the generator represents the rectifier system, L and C are the smoothing section, and R represents the resistance of the apparatus being supplied.

From *Graph 22* it is clear that the device must be operated well above the resonant frequency. If ω is the ripple frequency and $\omega_o = 1/(LC)^{1/2}$, as usual, then if ω/ω_o is less than unity, no filtering is achieved whatsoever, whilst if $\omega = \omega_o$ and the load resistance is large (Q high), far from attenuating the ripple the network may magnify it! In smoothing section design it is essential to check that the L - C product is adequate.

CLASSICAL FILTER

By inspection, we can write down that the transmission characteristic of the network is

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{\frac{R \frac{1}{j\omega C}}{R + \frac{1}{j\omega C}}}{j\omega L + \frac{R \frac{1}{j\omega C}}{R + \frac{1}{j\omega C}}}$$

which simplifies to

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{1}{1 + \frac{j\omega L}{R} - \omega^2 LC}$$

Now let us put in some typical values. If the apparatus consumes 100 milliamps at 350 V, then $R = 3,500$ ohms. C might well be 16 microfarads, and L 20 henries. If the rectification system is half-wave (see next chapter) the ripple frequency is 50 cycles and $\omega = 2\pi F$

Then
$$\omega^2 LC = \frac{4\pi^2 \times 50^2 \times 20 \times 16}{10^6} \approx 32$$

and
$$j\omega L/R = \frac{2\pi \times 50 \times 20}{3,500} \approx 1.8$$

Clearly the $\omega^2 LC$ term is much the most important, and for practical purposes it is sufficiently accurate to write

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{\omega^2 LC}$$

If the rectification is half-wave, this is not far from

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{10}{\text{Inductance in henries} \times \text{Capacitance in microfarads}}$$

and if the rectification is full-wave

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{2.5}{\text{Inductance in henries} \times \text{Capacitance in microfarads}}$$

CLASSICAL FILTER

In earlier chapters we have distinguished between the potential divider (whose behaviour is easy to compute if it is not required to feed a load, but which becomes altogether more involved when a load is connected) and the attenuator, which is intended for interposition between a matched generator and load, and whose component values are chosen so that matching conditions are preserved. We have also discussed some R-C filters (which are in effect frequency-conscious potential-dividers) and their performances have been worked out on the assumption that no load resistance is connected. In practice, this means the load resistance must be very high and preferably almost infinite—a valve.

The classical filter bears the same relationship to the R-C filter that the attenuator does to potential divider. That is, a classical filter is intended to work between a matched generator and load. Classical filters, like attenuators, are made up of sections which may be *T* or pi, and any number may be connected in cascade provided they are all designed for the correct characteristic resistance. The overall performance is then the sum of the several performances.

T-section

In *Figure 5.13* we see such a filter section in the *T* form, working between impedances Z_0 . For the present we work entirely in terms of impedance, in the interest of generality. The section comprises elements, as yet unspecified, of impedance $Z_1/2$ in the series arms and Z_2 in the shunt arm.

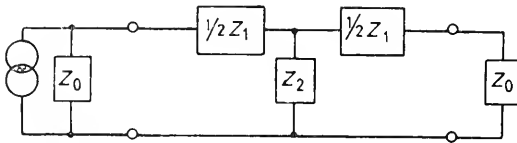


Figure 5.13

The requirement for matching conditions to be fulfilled is that on removing one of the terminating impedances Z_0 and looking into the filter, we see an impedance Z_0 , that is

$$Z_0 = \frac{Z_1}{2} + \frac{Z_2 \left(\frac{Z_1}{2} + Z_0 \right)}{Z_2 + \frac{Z_1}{2} + Z_0}$$

whence
$$Z_0^2 = Z_1 Z_2 \left(1 + \frac{Z_1}{4Z_2} \right)$$

If Z_1 and Z_2 are pure reactances, and the bracketed term is positive, Z_0 is pure resistive, for both the product and the ratio of *j* terms are real numbers.

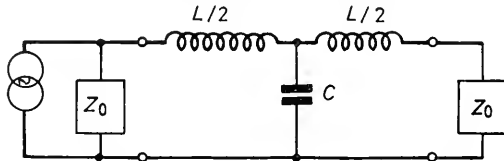


Figure 5.14

Low-pass filter

Value of elements—If the filter is to be low-pass we make Z_1 an inductance L and Z_2 a capacitance C (*Figure 5.14*). Then

$$Z_0^2 = j\omega L \left(\frac{-j}{\omega C} \right) \left(1 + \frac{j\omega L}{-4j/\omega C} \right)$$

$$Z_0 = \left(\frac{L}{C} \right)^{1/2} \left(1 - \frac{\omega^2 LC}{4} \right)^{1/2}$$

CLASSICAL FILTER

Clearly the arrangement is by no means perfect, for to preserve matching the generator and load impedances evidently have to alter in a special way with frequency, which is hardly practicable. When ω is small $Z_0 \doteq (L/C)^{1/2}$, a resistance; when ω equals $2/(LC)^{1/2}$, $Z_0 = 0$; and when $\omega > 2/(LC)^{1/2}$ the bracketed term becomes imaginary, which means Z_0 has turned into a reactance!

In practice the best we can do is to arrange for matching to be correct within the pass-band, which we shall show lies over frequencies below $\omega = 2/(LC)^{1/2}$. We let the load and generator impedances be constant resistances, equal to $(L/C)^{1/2}$ (Figure 5.15).

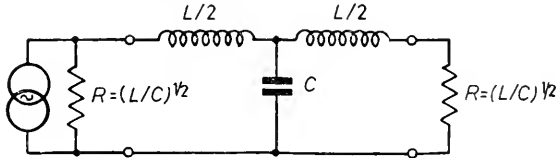


Figure 5.15

At $\omega = 2/(LC)^{1/2}$, Z_0 becomes zero because a series resonance occurs in the filter between C , and the two inductances $L/2$ in parallel (the net inductance is $L/4$, so the resonant frequency should be at $\omega = 1/(LC/4)^{1/2} = 2/(LC)^{1/2}$ which is right). This frequency is ω_C , the cut-off frequency for the filter; we dignify it with the name cut-off rather than turn-over because the transmission characteristic of L-C filters is much squarer than that of R-C filters, as we shall see.

Combining the equations $\omega_C = 2/(LC)^{1/2}$ and $R = (L/C)^{1/2}$ gives us

$$L = \frac{2R}{\omega_C} \quad \text{and} \quad C = \frac{2}{\omega_C R}$$

as expressions for the values of our filter elements. Let us see how they will behave.

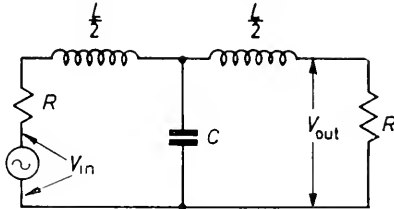


Figure 5.16

Transmission characteristic—In Figure 5.16, by inspection we have

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{R}{R + j\omega \frac{L}{2}} \cdot \frac{\left(R + j\omega \frac{L}{2} \right) \cdot -\frac{j}{\omega C}}{R + j\omega \frac{L}{2} - \frac{j}{\omega C}} \cdot \frac{\left(R + j\omega \frac{L}{2} \right) \cdot -\frac{j}{\omega C}}{R + j\omega \frac{L}{2} - \frac{j}{\omega C} + j\omega \frac{L}{2} + R}$$

which simplifies to

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{1}{2} \left\{ \frac{1}{\left(1 - \omega^2 \frac{L}{2} C\right) + j \left(\frac{\omega L}{2R} - \frac{\omega^3 L^2 C}{8R} + \frac{\omega C R}{2}\right)} \right\}$$

Now let $L/2 = R/\omega_C$ and $C = 2/\omega_C R$

$$\begin{aligned} \frac{V_{\text{out}}}{V_{\text{in}}} &= \frac{1}{2} \left\{ \frac{1}{\left[1 - 2\left(\frac{\omega}{\omega_C}\right)^2\right] + j \left[2\frac{\omega}{\omega_C} - \left(\frac{\omega}{\omega_C}\right)^3\right]} \right\} \\ \left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| &= \frac{1}{2} \left\{ \frac{1}{\left[1 - 2\left(\frac{\omega}{\omega_C}\right)^2\right]^2 + \left[2\frac{\omega}{\omega_C} - \left(\frac{\omega}{\omega_C}\right)^3\right]^2} \right\}^{1/2} \\ &= \frac{1}{2} \left(\frac{1}{\left\{1 + \left(\frac{\omega}{\omega_C}\right)^6\right\}^{1/2}} \right) \end{aligned}$$

The one-half here represents the usual loss of e.m.f. consequent upon the use of a matched load. The bracketed term describes the filter performance.

By the usual process the phase shift is seen to be

$$\phi = \tan^{-1} \frac{2\left(\frac{\omega}{\omega_C}\right) - \left(\frac{\omega}{\omega_C}\right)^3}{1 - 2\left(\frac{\omega}{\omega_C}\right)^2}$$

Extent to which matching is preserved—In Figure 5.17, the impedance

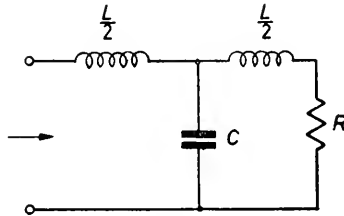


Figure 5.17

looking in at the front of the filter is

$$\begin{aligned} Z &= j\omega \frac{L}{2} + \frac{\left(R + j\omega \frac{L}{2}\right) \cdot \frac{1}{j\omega C}}{R + j\omega \frac{L}{2} + \frac{1}{j\omega C}} \\ &= \frac{\frac{L}{C} - \omega^2 \left(\frac{L}{2}\right)^2 + j\left(\omega \frac{L}{2} R - \frac{R}{\omega C}\right)}{R + j\left(\frac{\omega L}{2} - \frac{1}{\omega C}\right)} \end{aligned}$$

CLASSICAL FILTER

Putting in R/ω_C for $L/2$ and $2/\omega_C R$ for C

$$Z = \frac{R^2 - R^2 \left(\frac{\omega}{\omega_C}\right)^2 + j \left[\frac{\omega}{\omega_C} R^2 - \frac{R^2}{2} \left(\frac{\omega_C}{\omega}\right) \right]}{R + j \left[R \frac{\omega}{\omega_C} - \frac{R}{2} \frac{\omega_C}{\omega} \right]}$$

which simplifies eventually to

$$Z = R \left\{ \frac{\left(\frac{\omega_C}{2\omega}\right)^2 + j \left[\left(\frac{\omega}{\omega_C}\right)^3 - \frac{\omega}{2\omega_C} \right]}{\left(\frac{\omega}{\omega_C}\right)^2 + \left(\frac{\omega_C}{2\omega}\right)^2} \right\}$$

In *Graphs 26, 27 and 28* the transmission characteristic, phase shift and terminal impedances for a classical low-pass filter section are plotted. The transmission characteristic is clearly much superior to an R-C filter section, since it possesses an attenuation slope of 18, instead of 6, dB's per octave. Furthermore, it is better than a 3-stage R-C filter of tapered sections, for whereas the latter would be 9 dB's down at the cut-off frequency, the L-C filter is only 3 dB's down. The characteristic is therefore squarer, more like the ideal filter, than anything that can be done with passive-element filters comprising R's and C's.

The impedance presented to the generator and load is substantially resistive and equal to R in the pass-band.

High-pass filter

In the classical high-pass section (*Figure 5.18*) the series arms are capaci-

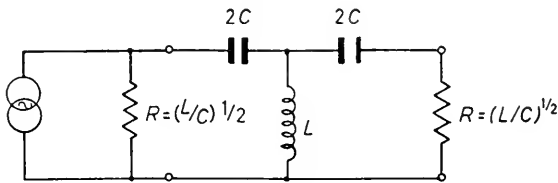


Figure 5.18

tances, and the shunt arm an inductance. We have, as before

$$Z_0^2 = Z_1 Z_2 \left(1 + \frac{Z_1}{4Z_2} \right)$$

but now, $Z_1 = -j/\omega C$ and $Z_2 = j\omega L$. So $Z_1 Z_2$ is still L/C , but $Z_1/4Z_2 = -1/(4\omega^2 LC)$. At the cut-off frequency, series resonance occurs in the filter and Z_0 goes to zero. $(1 - \{1/(4\omega_C^2 LC)\})$ is therefore also equal to zero, hence $\omega_C = 1/\{2(LC)^{1/2}\}$. In the pass-band, $Z_0^2 \cong L/C$, $R = (L/C)^{1/2}$.

INDUCTANCES, CAPACITANCES AND RESISTANCES

Combining the equations $R = (L/C)^{1/2}$ and $\omega_C = 1/\{2(LC)^{1/2}\}$ we get for the values of the filter elements

$$L = \frac{R}{2\omega_C} \quad \text{and} \quad C = \frac{1}{2\omega_C R}$$

We find the transmission characteristic, phase shift and terminal impedance in a manner similar to that for the low-pass section. The transmission factor comes to

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{2} \left\{ \frac{1}{\left[1 + \left(\frac{\omega_C}{\omega} \right)^6 \right]^{1/2}} \right\}$$

and the phase shift to

$$\phi = \tan^{-1} \frac{2 \frac{\omega_C}{\omega} - \left(\frac{\omega_C}{\omega} \right)^3}{1 - 2 \left(\frac{\omega_C}{\omega} \right)^2}$$

and the impedance seen by generator and load is

$$R \left\{ \frac{\left(\frac{\omega}{2\omega_C} \right)^2 - j \left[\left(\frac{\omega_C}{\omega} \right)^3 - \frac{\omega_C}{2\omega} \right]}{\left(\frac{\omega_C}{\omega} \right)^2 + \left(\frac{\omega}{2\omega_C} \right)^2} \right\}$$

These are also plotted in *Graphs 26, 27 and 28*, and reveal the symmetrical properties of the two filters.

'Pi' section filter

Suppose we take a string of similar *T* sections (*Figure 5.19*), then this may be redrawn as in *Figure 5.20*, and redrawn again as in *Figure 5.21*. Then, between each pair of dotted lines, we have a pi section.

Low-pass—Clearly the low-pass *T* in *Figure 5.22* transforms to the pi version as in *Figure 5.23*. The two networks are equivalent up to a point; they have the same transmission characteristic and phase shift; but the impedance seen looking into the end of the pi section is

$$Z = R \left\{ \frac{\left(\frac{\omega_C}{2\omega} \right)^2 - j \left[\left(\frac{\omega}{\omega_C} \right)^3 - \left(\frac{\omega}{2\omega_C} \right) \right]}{\left(\frac{\omega}{\omega_C} \right)^4 - \left(\frac{\omega}{\omega_C} \right)^2 + \left(\frac{\omega_C}{2\omega} \right)^2} \right\}$$

High-pass—By a similar argument the high-pass *T* section in *Figure 5.24* is, in the pi form, given by *Figure 5.25*. Once again the two have similar

CLASSICAL FILTER

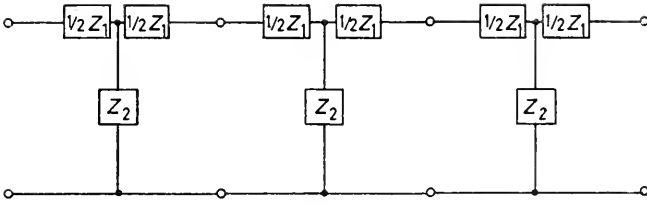


Figure 5.19

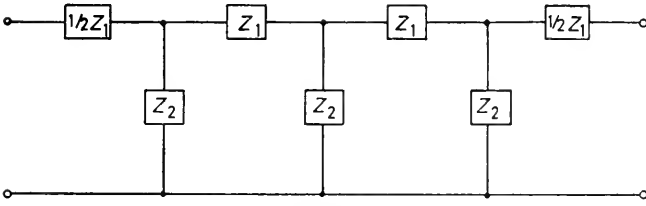


Figure 5.20

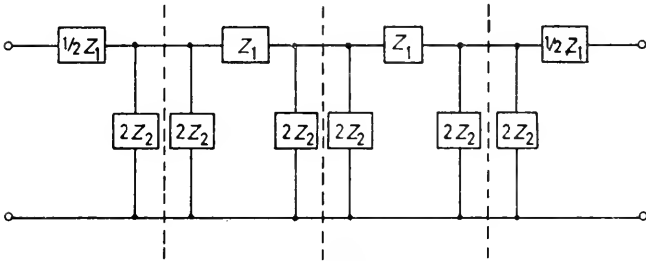


Figure 5.21

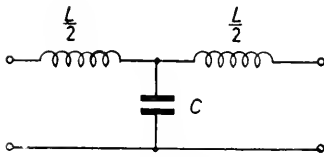


Figure 5.22

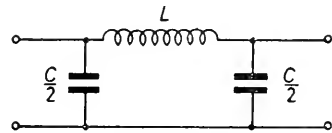


Figure 5.23

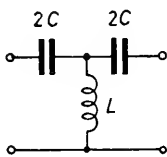


Figure 5.24

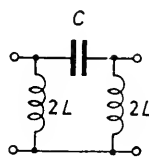


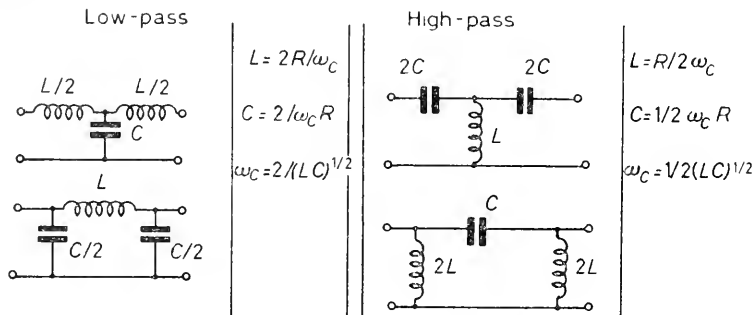
Figure 5.25

INDUCTANCES, CAPACITANCES AND RESISTANCES

transmission and phase characteristics but the terminal impedance of the pi version is

$$Z = R \left\{ \frac{\left(\frac{\omega}{2\omega_C}\right)^2 + j\left[\left(\frac{\omega_C}{\omega}\right)^3 - \left(\frac{\omega_C}{2\omega}\right)\right]}{\left(\frac{\omega_C}{\omega}\right)^4 - \left(\frac{\omega_C}{\omega}\right)^2 + \left(\frac{\omega}{2\omega_C}\right)^2} \right\}$$

Summary



OTHER L-C FILTERS

The subject of filters is a vast one, and we can do no more than mention some of the other types. Readers faced with particularly stringent filter problems should refer to a specialist work on the subject (E. A. Guillemin, *Communication Networks*, New York; Wiley).

The transmission characteristic of the classical filter is of the type known as Butterworth or maximally flat. It is possible to obtain a squarer characteristic by using a Tchebycheff type of filter, whose response is typically of the form shown in *Figure 5.26*; evidently the transition between the horizontal

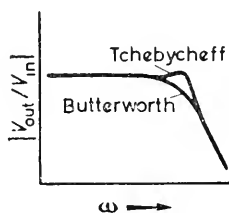


Figure 5.26

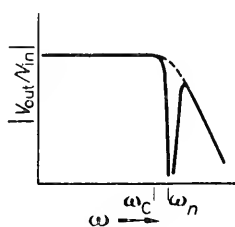


Figure 5.27

and sloping parts of the curve is more abrupt. Tchebycheff filter sections are similar in configuration to the Butterworth type, but the component values are different.

If the filter has to discriminate between two frequencies rather close together, a useful device is the *m*-derived section, which has—in the low-pass version—a characteristic like *Figure 5.27*. This is a Butterworth response with a bottomless trough in the transition region; that is, it is a null-transmission network. If the frequency of null-transmission is to be ω_n , then to get an

OTHER L-C FILTERS

m -derived filter we take a classical filter of cut-off frequency ω_C and define m so that

$$m = \left\{ 1 - \left(\frac{\omega_C}{\omega_n} \right)^2 \right\}^{1/2} \quad (\text{low-pass filters})$$

and
$$m = \left\{ 1 - \left(\frac{\omega_n}{\omega_C} \right)^2 \right\}^{1/2} \quad (\text{high-pass filters})$$

Thus, m varies in the range 0 to 1 and is small when ω_n and ω_C are close together.

Then a low-pass T section transforms to m -derived by rebuilding it either as in *Figure 5.28b* or as in *Figure 5.28c*. In the first of these null-transmission

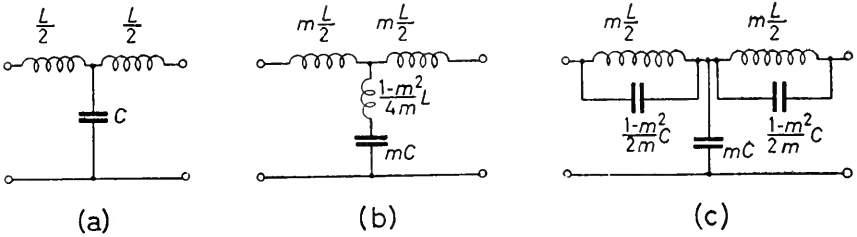


Figure 5.28

is obtained by series resonance in the shunt arm, and in the second, by a parallel resonance in each series arm.

Similarly a high-pass T section, in the m -derived form, is either as in *Figure 5.29b*, or as in *Figure 5.29c*. Once again, null-transmission occurs as a result of resonance in the appropriate arm.

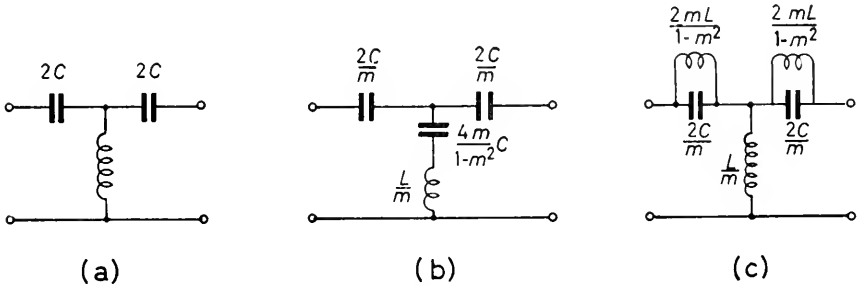


Figure 5.29

In deciding which of the two alternative m -derived versions of each filter type to use in a particular case, bear in mind that in general, capacitors are more readily available than inductors, and cheaper; work out the values required for each alternative, and see which is the more convenient.

If m can be arranged to be in the neighbourhood of 0.6, it can be shown that the impedance looking in at either end of the filter remains near to R throughout the pass-band and to a frequency very close to ω_C (*Figure 5.30*).

The m -derived pi sections may be found by considering a long chain of

INDUCTANCES, CAPACITANCES AND RESISTANCES

m -derived T sections, which is then divided up to form pi's, in the same manner as classical pi sections may be found from classical T sections.

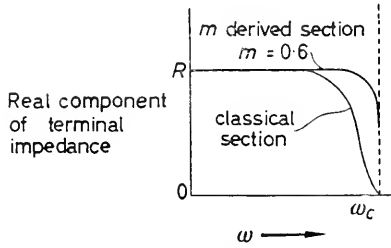


Figure 5.30

The band-pass filter—In order to pass a band of frequencies between ω_1 —the lower cut-off frequency—and ω_2 —the upper cut-off frequency—we could of course connect in cascade a high-pass filter section cutting off at ω_1 , and a low-pass section cutting off at ω_2 , the sections both being designed for the appropriate characteristic resistance R .

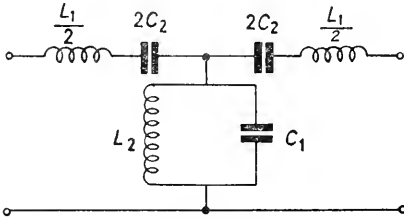


Figure 5.31

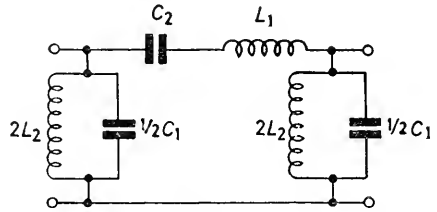


Figure 5.32

There is, however, a special band-pass section, which is shown in Figure 5.31 in the T form, and in Figure 5.32 in the pi form. The design equations are

$$L_1 = \frac{2R}{\omega_2 - \omega_1}$$

$$L_2 = \frac{(\omega_2 - \omega_1)R}{2\omega_1\omega_2}$$

$$C_1 = \frac{2}{(\omega_2 - \omega_1)R}$$

$$C_2 = \frac{\omega_2 - \omega_1}{2\omega_1\omega_2 R}$$

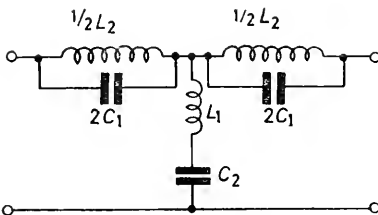


Figure 5.33

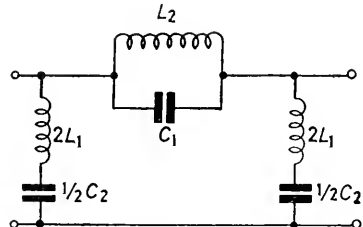


Figure 5.34

The band-stop filter—This is a section having the inverse property to the above, in that it passes all frequencies *except* those lying between a lower, ω_1 , and an upper, ω_2 . In the T form it is as in Figure 5.33, and in the pi form as in Figure 5.34.

The design equations are

$$L_1 = \frac{R}{2(\omega_2 - \omega_1)} \quad L_2 = \frac{2(\omega_2 - \omega_1)R}{\omega_1\omega_2}$$

$$C_1 = \frac{1}{2(\omega_2 - \omega_1)R} \quad C_2 = \frac{2(\omega_2 - \omega_1)}{\omega_1\omega_2 R}$$

Filters to work between unequal resistances

We have dealt at some length with the classical high- and low-pass filters, and have mentioned briefly the band-pass, band-stop, and m -derived filters. All these are supposed to work between equal generator and load resistances, and the reasons for making this provision are: (1) the case is theoretically important because it is the condition for maximum power transfer; and (2) it enables a filter of complicated transmission characteristic to be assembled from simple sections, the overall performance being merely the sum of the several performances (on a dB scale).

It often happens in electronics that filters are required for parts of a circuit where maximum power transfer conditions do not obtain. In valve circuits we are usually more interested in voltage than in power, and maximum voltage transfer occurs when the load resistance is much larger than the generator resistance. Further, such filters are often of a simple kind, requiring only one section. It is therefore pertinent to inquire whether filters can be built when the resistances between which they are to work are different. The answer is that they can, and we shall indicate the general procedure and illustrate with an example. We restrict ourselves to sections having a Butterworth response, i.e. whose transmission characteristic is similar to that for a classical filter.

For a Butterworth response, if the filter possesses n reactive elements, the transmission characteristic is 3 dB's down at the cut-off frequency and is asymptotic to a slope of $6n$ dB's per octave, and

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{\left\{ 1 + \left(\frac{\omega}{\omega_C} \right)^{2n} \right\}^{1/2}} \quad (\text{high-pass}) \quad \text{or} \quad \frac{1}{\left\{ 1 + \left(\frac{\omega}{\omega_C} \right)^{2n} \right\}^{1/2}} \quad (\text{low-pass})$$

In our case, $n = 3$, so for a low-pass filter

$$\left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| = \frac{1}{\left\{ 1 + \left(\frac{\omega}{\omega_C} \right)^6 \right\}^{1/2}}$$

Looking back to the calculations for the transmission characteristic of the classical low-pass filter, we see that the penultimate stage is of the form

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{1}{(1 + A\omega^2 + B\omega^4 + D\omega^6)^{1/2}}$$

and happily A and B turn out to be zero, and D to be $(1/\omega_C)$. This happens because of the particular values we gave to the reactive elements of the section, in terms of ω_C and R . Had the filter elements or the R been different,

INDUCTANCES, CAPACITANCES AND RESISTANCES

A and B would not have been zero and the response of the filter would not have been Butterworth.

If the R values are altered, how can we make the response return to Butterworth? By finding new values of L and C such that A and B are zero and $D = (1/\omega_c)^6$. We take as an example a very important case in valve work, a filter to operate between a generator of resistance R and a load of infinity. We take the low-pass case.

Referring to *Figure 5.35*, the first point to notice is that the section must

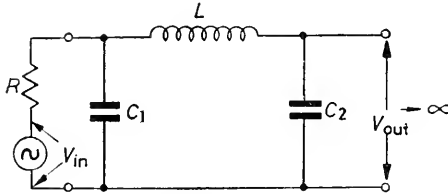


Figure 5.35

be of the pi form. If it were a T , since the load current is zero, the right-hand horizontal element would also pass no current; if it passes no current it can make no contribution to the working of the filter. The filter would degenerate to a two-element one, producing an attenuation slope of only 12 dB/octave.

Proceeding as usual we see by inspection that

$$\frac{V_{out}}{V_{in}} = \frac{1}{j\omega C_2} \cdot \frac{\frac{1}{j\omega C_1} \left(\frac{1}{j\omega C_2} + j\omega L \right)}{\frac{1}{j\omega C_1} + \frac{1}{j\omega C_2} + j\omega L} \cdot \frac{1}{\frac{1}{j\omega C_1} \left(\frac{1}{j\omega C_2} + j\omega L \right) + R} \cdot \frac{1}{\frac{1}{j\omega C_1} + \frac{1}{j\omega C_2} + j\omega L}$$

which simplifies to

$$\frac{-1}{(\omega^2 LC - 1) + j[\omega^3 C_1 C_2 R - \omega R(C_1 + C_2)]}$$

Hence,

$$\left| \frac{V_{out}}{V_{in}} \right| = \frac{1}{\{1 + \omega^2[R^2(C_1 + C_2)^2 - 2LC_2] + \omega^4[L^2C_2^2 - 2R^2LC_1C_2(C_1 + C_2)] + \omega^6[C_1^2C_2^2R^2L^2]\}^{1/2}}$$

If this is to be Butterworth, then:

(1) The coefficient of ω^6 must be

$$\left(\frac{1}{\omega_c} \right)^6 \quad \therefore (\omega_c)^6 = \frac{1}{C_1^2 C_2^2 R^2 L^2} \quad \text{and} \quad \therefore \omega_c^3 = \frac{1}{C_1 C_2 R L}$$

(2) The coefficient of ω^4 must be zero. Therefore

$$L^2 C_2^2 = 2R^2 L C_1 C_2 (C_1 + C_2)$$

OTHER L-C FILTERS

or

$$LC_2 = 2R^2C_1(C_1 + C_2)$$

(3) The coefficient of ω^2 must be zero. Therefore

$$2LC_2 = R^2(C_1 + C_2)^2$$

Combining (2) and (3) it emerges that $C_2 = 3C_1$ (i)

Putting this result back in (3), we find $L = \frac{8}{3} C_1 R^2$ (ii)

Using (i) and (ii) in (1) we get $\omega_C = \frac{3}{2RC_2}$ (iii)

These are not quite in the form required for design. Re-arranging (iii)

$$C_2 = \frac{3}{2R\omega_C} \quad \text{is one design equation} \quad \text{(a)}$$

from (i)

$$C_1 = \frac{1}{2R\omega_C} \quad \text{is another} \quad \text{(b)}$$

Substituting (b) in (ii) gives

$$L = \frac{4R}{3\omega_C} \quad \text{as the third} \quad \text{(c)}$$

Combining (a) and (c), it emerges that the section cuts off at

$$\omega_C = \left(\frac{2}{LC_2} \right)^{1/2} .$$

DIODE CIRCUITS

A diode* is a device which exhibits a markedly lower resistance to the passage of electric current in one direction than in the other (*Figure 6.1*). If the current passed by a diode is plotted as a function of the potential difference across the terminals, the graph is found to comprise two more or less straight lines possessing different slopes, and the slopes of these lines give the forward resistance and backward resistance of the diode. The figure of merit for diodes is the ratio of these slopes, known colloquially as the 'front-to-back ratio'.

If the diode is thermionic, its back resistance is effectively infinite and the arrangement is therefore a perfect valve (*Figure 6.2*). Non-thermionic diodes



Figure 6.1

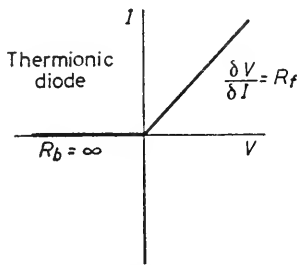


Figure 6.2

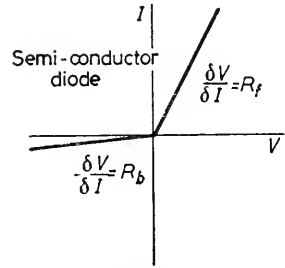


Figure 6.3

make use of semi-conducting materials such as copper oxide, selenium, germanium and silicon. With these the backward resistance is regrettably finite, practical front-to-back ratios being upwards of 100 (*Figure 6.3*). In this chapter we enumerate and discuss some of the functions of diodes.

RECTIFICATION

Apart from valve heaters, electronic apparatus is powered largely by direct voltages. Where apparatus is to be run from the supply mains, nowadays mainly alternating voltage, the necessary conversion is done by *power rectification*. Again, it often happens that a signal is represented by the amplitude of an alternating voltage and we wish for some reason to convert the signal to a direct voltage. Here again we use the rectification process, this time *signal rectification*. Signal rectifiers are also sometimes called detectors, not a good name, but one rooted in the history of wireless telegraphy. In brief,

* Strictly, a diode is any two-electrode device, but here we restrict the term to those in which use is made of asymmetric resistance properties.

Certain devices sometimes called diodes—notably the cold cathode soft diodes—are so named in the sense of having two electrodes, and are not relevant here.

RECTIFICATION

then, rectification is the process of producing from alternating supplies an output which is direct, proportional to the input and which may or may not be constant with time. If it is steady with time it is called smooth.

Half-wave rectifier—Suppose a diode, load resistance and generator are connected in series as shown in *Figure 6.4*. Before describing what happens

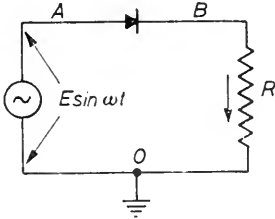


Figure 6.4

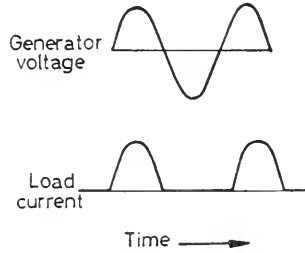


Figure 6.5

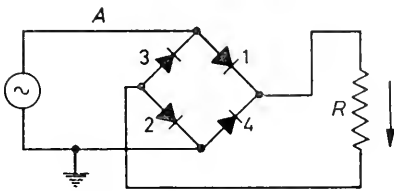
we introduce a convention, universal in electronics, which greatly facilitates discussion, the earth connection at *O*. It does not greatly matter whether *O* is actually earthed or not but it establishes a datum potential for *O* and we can then speak simply of the potential at *A* being positive, or negative, or rising, i.e. becoming more positive, or falling, instead of saying the *difference* of potential between *A* and *O* does so and so.

We have then, due to the generator, the potential at *A* swinging up and down sinusoidally between $+E$ and $-E$. When *A* is positive the diode exhibits very low resistance and a current

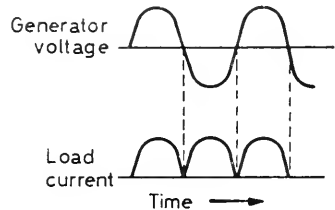
$$\left[\frac{E \sin \omega t}{R} \right]$$

flows through the load. When *A* is negative the diode resistance is high and negligible current flows. This is illustrated in *Figure 6.5* and is simple half-wave rectification. The load current is uni-directional but intermittent.

Full-wave rectifier—A steadier output current is had by a rectifier bridge in which four diodes are arranged as in *Figure 6.6a*. When *A* is positive



(a)



(b)

Figure 6.6

diodes 1 and 2 conduct; 3 and 4 are in a high resistance condition or 'cut off'; power is delivered to the load. When *A* is negative, the situation reverses as regards the diodes but power is still delivered to the load (*Figure 6.6b*). This

DIODE CIRCUITS

arrangement is called full-wave rectification. It uses rather a lot of diodes and it is possible to do better than this if the generator is of a special kind, that is, it has a 'centre-tap'. In *Figure 6.7* the generator seen by the diodes is a

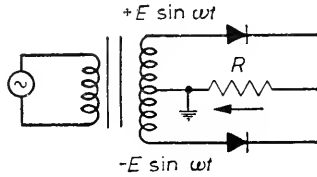


Figure 6.7

transformer with a centre-tapped secondary winding. Then we have two half-wave circuits working back to back, alternately, and the load current is of the same form as that for the bridge.

This type of circuit is the most commonly used for power rectification in electronics and the unqualified phrase 'full-wave rectifier' circuit refers to this rather than to the bridge arrangement. It is open to the objection that the transformer insulation has to be able to withstand twice the voltage it would have to in the bridge configuration. The pulsatory outputs of these three rectifier circuits are of very little use as they stand. They are all right for accumulator charging, but mostly are more useful if they are smoothed. We should, however, comment on one extremely important case of unsmoothed bridge rectification.

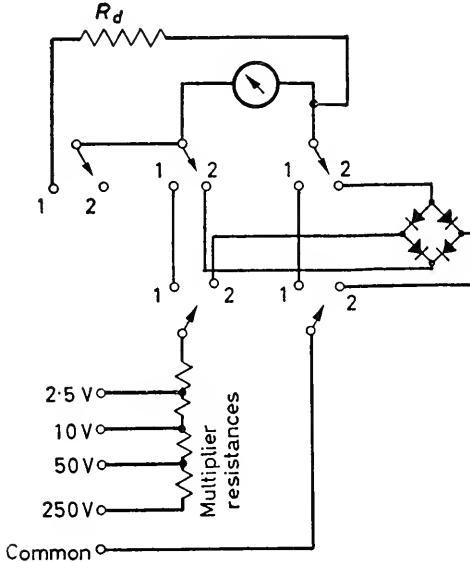


Figure 6.8

The multi-range meter—These versatile instruments are much in evidence in electronic work and it is necessary to issue a word of warning about their use. The voltmeter part of one of these might be arranged like *Figure 6.8*.

SMOOTHED RECTIFICATION

When a direct voltage is to be read, all the switches are in position 1 and the circuit is by way of the appropriate multiplier resistance, straight to the moving coil unit, which has the resistance R_d connected across it. When the instrument is switched to read alternating voltage, all contacts are altered to position 2, and the circuit reduces to *Figure 6.9*. When any alternating

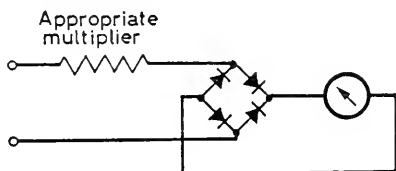


Figure 6.9

voltage of any waveform is connected, the deflection of the pointer is proportional to the average of the rectified current. If the input voltage is sinusoidal, the average value is

$$\frac{1}{\pi} \int_0^{\pi/\omega} V \cdot \sin \omega t \cdot dt = \frac{2V}{\pi} = 0.637 V$$

and the pointer deflection will be proportional to this. The reading which is required on the scale is the R.M.S. value of V , which is $0.707V$. Thus, in order that the instrument shall read correctly on alternating voltage ranges it has to have its sensitivity increased by a factor $0.707/0.637 = 1.11$, and this is achieved by removing automatically from the circuit the diverter resistance R_d . The reading is then right for a pure sinusoidal alternating voltage input but not for any other kind of waveform. Thus, a waveform which often occurs in electronics is the so-called square wave (*Figure 6.10*) which has the

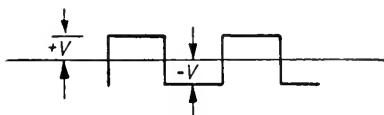


Figure 6.10

property that its R.M.S. and average after rectification are the same. If this is of amplitude V , the reading indicated by the pointer will be 11.1 per cent high. If the meter is of the kind which contains a transformer it may be worse still. Much could be written about the interpretation of alternating voltage voltmeter readings but in the author's opinion if the waveform is at all peculiar it is best to look at it on a measuring oscilloscope; for special waveforms it is in any case a moot point how useful or significant an average reading is.

SMOOTHED RECTIFICATION

To smooth the output of a rectifying circuit we can either: (1) put an inductance in series with the load; or (2) put a capacitance in parallel with the load. The first method is called the choke input circuit, while the second is the capacitor input circuit.

DIODE CIRCUITS

Capacitor input rectifying circuit

This is shown in the half-wave form in *Figure 6.11*. It appears in apparatus of the a.c./d.c. type for power rectification (because no transformer can be used, so no centre-tap is available) and is the most usual form for signal rectification. A full-wave arrangement is shown in *Figure 6.12* and is almost

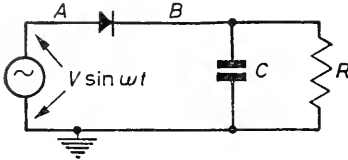


Figure 6.11

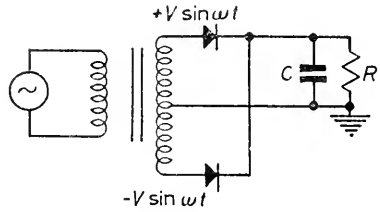


Figure 6.12

universal for supplying power to an electronic apparatus where the current is less than, say, 250 milliamps. Very few pieces of equipment consume more than this so that this circuit may be regarded practically as standard.

To see how these circuits work, consider the half-wave case. Suppose R , the load, be temporarily disconnected and C discharged and the generator then switched on. When A swings positive the diode conducts, B is carried positive too and C is charged from the generator. When A and B reach the potential V , A begins to move negative again, reversing the potential difference across the diode, whose resistance therefore becomes high. The current cannot now flow into or out of the capacitance C , which remains charged to a potential V indefinitely. The important fact thus emerges that the no-load voltage of a capacitance-input rectifying circuit is equal to the peak generator voltage.

Now let the load be reconnected and suppose the potential at A is just beginning to go negative (point a in *Figure 6.13*). The potential difference

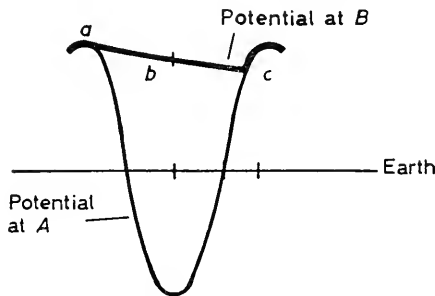


Figure 6.13

across the diode reverses, switching it to high resistance, so that the charged capacitance is left connected only to the load, to which it delivers load current. The voltage across the capacitance falls in an approximately linear manner with time (b in *Figure 6.13*) until point c is reached when A once more rises above B , switching the diode to low resistance and carrying B back up to V . This rhythmic fluctuation in the output voltage is called ripple.

SMOOTHED RECTIFICATION

If the load is made heavier, R is reduced, C has to supply more current and discharges further between charging periods. Thus we see that on increasing the load on a capacitor input rectifier circuit the ripple voltage increases and the mean output voltage falls. The latter is in fact $V - \frac{1}{2}$ (peak to peak ripple) (*Figure 6.14*).

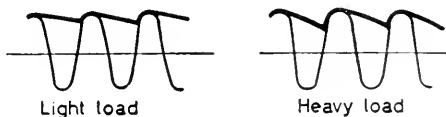


Figure 6.14

Application of this circuit to power supply—The regulation of a power supply is said to be good when the output voltage falls only a little with increase of load current. On the basis of what we have said so far it looks as if ripple may be reduced and regulation improved as much as we like by increasing the value of the capacitance. In point of fact the ripple is reduced but the regulation may get worse and the diode even be destroyed. Let us see why this is.

Referring to *Figure 6.15* it is clear that diode current only passes during

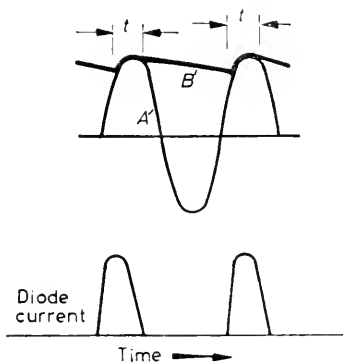


Figure 6.15

the short period t . Since all the output current must at some time have passed through the diode, it follows that the diode current must be of a magnitude many times the load current. Further, it is clear that if the capacitance be increased for a given load, the diode conduction time becomes even shorter and the peak diode current must of necessity become larger in order that power may be supplied to the load at the same rate. It is these large peak currents which are damaging to thermionic diodes, though the makers of selenium diodes claim that they are not harmed by large instantaneous currents.

When we take into account the resistance of the generator and the forward resistance of the diode we see why the regulation suffers. The high peak currents passed by the diode produce unexpectedly large voltage drops in the generator internal resistance and the diode at the instant of conduction, so that point B never reaches E , only something lower.

DIODE CIRCUITS

From the point of view of regulation, then, there is an optimum value of C ; the calculations involved in finding it are complicated, but fortunately the manufacturers of rectifying diodes are very helpful, in their published data, with recommendations on this point. Practical values range between 4 and $32 \mu\text{F}$.

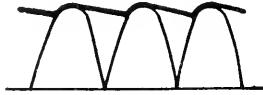


Figure 6.16

It is now evident why the full-wave system of power rectification is so much to be preferred. Diode conduction times occur twice as often (*Figure 6.16*) so for a given load and capacitance the peak diode current is halved, thus regulation is improved. Alternatively, more capacitance can be used for the same regulation, thus reducing the ripple. Another important point is that the fundamental ripple frequency is doubled, which facilitates further smoothing.

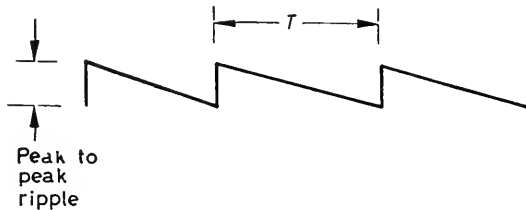


Figure 6.17

The ripple voltage may be estimated in the following way. Suppose all the diode current per cycle flows instantaneously, then the output waveform is triangular and the capacitance supplies load current for time T (*Figure 6.17*). The rate of discharge for the capacitance is constant for all reasonable values of C and equal to I/C volts per second where I is the load current. For half-wave systems on 50 cycle mains, $T = 0.02$ seconds, so the peak to peak ripple voltage is $0.02 I/C$. For a full-wave system $T = 0.01$ seconds so the peak to peak ripple is $0.01 I/C$.

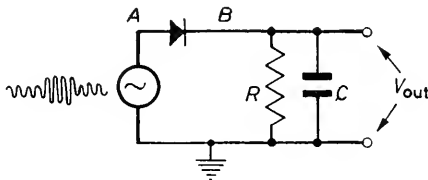


Figure 6.18

Application to signal rectification (Figure 6.18)—For signal rectification the load is usually constant and so regulation is not of any importance. Further, power levels are low and there is no question of destroying the diode. We are therefore free to vary C and see how doing this affects the

SMOOTHED RECTIFICATION

performance of the circuit in its ability to fulfill its task, namely, to produce a direct voltage which follows faithfully the amplitude of an alternating voltage. *Figure 6.19b*, a carrier wave, has been modulated by the pulse in *Figure 6.19a* and this modulated carrier is supposed to be the output of the generator.

When the carrier peak voltage rises from V_1 to V_2 , provided the generator and diode forward resistances are low, V_{out} reaches V_2 at the peak of the next positive excursion of point A , that is to say the response time cannot

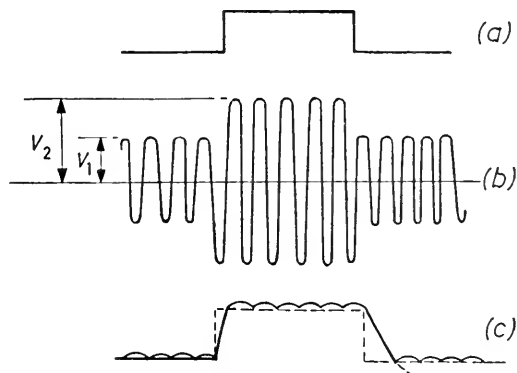


Figure 6.19

exceed the time of one cycle of the carrier (*Figure 6.19c*). At the end of the pulse matters are rather different, for now the potential at A never reaches that of B and the diode is continuously cut off until V_{out} falls once more to V_1 . This it does by approaching zero with a time constant CR until such time as its decay is arrested by renewed conduction in the diode.

Two points emerge: (1) distortion is more serious at the end of the pulse; (2) distortion at the end of the pulse is lessened when CR is small. R cannot be made small if the generator and diode forward resistances are to remain relatively negligible, therefore the reduction must be made in C .

The magnitude of the ripple in the output depends on the carrier frequency, the load and the value of C . For small ripple we need a high carrier frequency (there is then less time for the charge on C to leak away between cycles) or a large R (which we have already got) and a large C .

Requirements for low distortion at the pulse end, small c , and low distortion during the pulse, large C , are therefore conflicting. Signal rectifier circuits are quite easy to design when the carrier frequency is much higher than the modulation frequency. The procedure is to arrange that CR is much greater than the time of the cycle of carrier but much less than the time of the cycle of modulation: the geometric mean would be suitable. When the two frequencies are closer, the carrier being perhaps only ten times the highest modulation frequency, it is best to have a low C , which preserves the back edge of the pulse, and remove the ripple by subsequent low-pass filtering.

When the generator and diode forward resistance are not negligible compared with the load, the effect is twofold: (1) the rising phase of V_{out} will also be slowed up—several cycles of carrier at the new value of V_2 may be required before V_{out} reaches its final value; (2) V_{out} will not reach V_2 but

DIODE CIRCUITS

will level off at something lower. The *detector efficiency* is defined as the output voltage divided by the peak carrier voltage required to produce it. It is not a good term because it implies that signal rectifier circuits with poor detector efficiencies must be somehow unsatisfactory. If the efficiency is low because a bad diode has been used having a high forward resistance this is perfectly true. If it is low because a small value of C has been chosen in the interests of faithful response, then it is not.

INDUCTANCE-INPUT (OR CHOKE INPUT) RECTIFYING CIRCUIT

This is used for power rectification where the load current is so high that the capacitance input circuit would have excessive ripple and exceedingly poor regulation. It is always used in a full-wave system (*Figure 6.20*). The

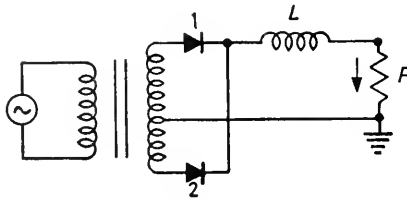


Figure 6.20

inductance and the load resistance may be regarded as forming a low-pass filter which will allow the steady direct-voltage component of the potential at $A = (2/\pi) V$ (as we saw in the section on multi-range meters) to pass on to the load R_L at B , but which will attenuate the ripple frequency component, so that the load current will be relatively steady. The cut-off frequency of this arrangement will be at $\omega_c = R/L$, and as for the best attenuation of the ripple frequency, ω_c must be much lower than the ripple frequency, we have that, for good smoothing, L must be large or R must be small. It follows that *choke input smoothing works best when the load is heavy* (contrast this with capacitance-input filtering). It can be shown that for a 50 cycle supply the load ripple current will not exceed 10 per cent of the load steady current if L in henries is greater than or equal to $R/470$, where R is in ohms. This is the 'optimum inductance'.

If the inductance is sufficient to satisfy a certain criterion of smoothing when the load is light, then it will be unnecessarily large when the load is heavy. In order to achieve greater economy in material here a special type of inductor known as a 'swinging choke' is commonly used, which makes use of saturation effects in the core. When the load is light, the current is small, the flux is small, the permeability of the iron is maximal and so is the inductance. As the current rises the iron partially saturates and the inductance falls, but the ratio L/R is approximately maintained. If a normal inductor were used which would maintain its inductance value up to the highest load current, a great deal of iron would be necessary and a very bulky component would be required.

The regulation of choke input rectifying circuits is very good for the

INDUCTANCE-INPUT RECTIFYING CIRCUIT

following reason. In *Figure 6.21* we have *Figure 6.20* again, but an additional capacitance C , usually present in this type of circuit in the form of extra smoothing, is connected across R and stabilizes the potential difference across R_L at $(2/\pi)V$. At first sight it would appear that the conduction time of diode (1) is that for which the potential at A exceeds $(2/\pi)V$, and

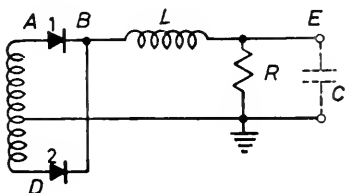


Figure 6.21

similarly for diode (2) and the potential at D (*Figure 6.22*): in fact it is even longer. Suppose point A is positive to earth but has passed its peak and is negative going: diode (1) is conducting and diode (2) is cut off. As A is falling there is a tendency for the inductance current to fall, which the inductance combats by generating a back e.m.f. across itself, the sense of which is to make B more negative than E . If the potential at E is stabilized, the effect of the back e.m.f. is to depress the potential at B , so that A can go further negative before diode 1 is cut off. Thus the conduction time of diode 1 is extended. The benefit conferred by the long conduction time is very low

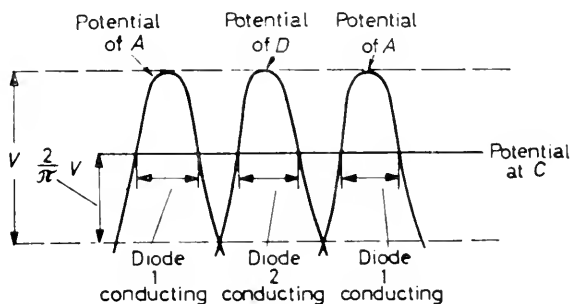


Figure 6.22

peak diode current and so low voltage drops. The same applies in the case of diode 2 and hence the regulation of the circuit is good.

There is a difficulty with choke-input circuits which requires bearing in mind. Referring to *Figure 6.21*, if the load is accidentally removed the inductance current is drastically reduced and so, therefore, is the contribution of the inductance to the operation of the circuit. When this happens the arrangement reverts to the capacitance-input type of rectifier and the output voltage rises from its normal value of about $2V/\pi$ to V . This may cause serious damage, and to prevent it happening it is necessary to ensure that some load is always present. Such a load is called a 'bleeder' (R_b in *Figure 6.23*). It can be shown that if for a 50 cycle supply $L \geq R_{Lmax}/940$ the output voltage will not rise above its normal value. This value of L is known as the 'critical inductance'. Design procedure is therefore as follows: given

DIODE CIRCUITS

that it is desired to rectify a certain voltage V , and current, $0-I_{\max}$, we have: minimum load resistance $R_{L\min}$ which will be connected $= V/I_{\max}$. Work out the optimum inductance from $L = R_{L\min}/470$, finding a swinging choke whose inductance is this at maximum current I_{\max} .

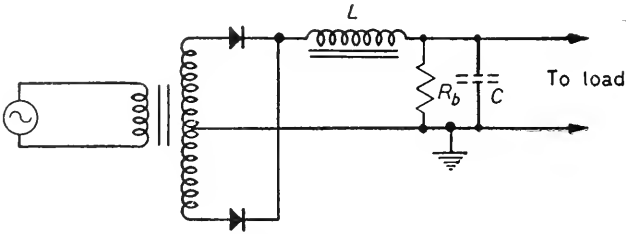


Figure 6.23

Find the inductance of the swinging choke at low values of current from the makers' characteristics or by experiment, and call this L' .

Work out the maximum permissible bleeder resistance R_b , from $L' = R_b/940$.

VOLTAGE MULTIPLIER RECTIFYING CIRCUITS

A number of pieces of apparatus in electronics, such as Geiger tubes and cathode ray tubes, require supplies of very low direct current and rather high voltage, say more than 1,500 volts. It is possible to derive such supplies by conventional rectification of the output of a special high-voltage secondary winding on the power transformer, but often inconvenient. If lower alternating voltage supplies are available it is possible to generate the necessary high-voltage required by special 'voltage multiplier' circuits. Since the currents to be supplied are small these circuits are elaborations of the simple capacitance-input half-wave circuit already described, and their regulation is therefore poor. Somewhat simplified descriptions of these circuits and their operation will now be given.

Symmetrical voltage doubler (Figure 6.24)—When terminal A of the generator is positive with respect to terminal B current flows in the upper part of

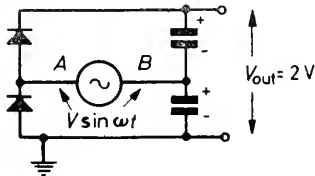


Figure 6.24

the circuit, charging the upper capacitance to a voltage V , with the polarity shown. When terminal A is negative with respect to terminal B the lower part of the circuit is operative and the lower capacitance is charged, also to a voltage V and with the polarity shown. Thus the output is $2V$.

The difficulty with this circuit is that both generator terminals are involved in the circuit action, which means that the generator is not generally available for supplying other rectifying circuits as well. By a modification of the

VOLTAGE MULTIPLIER RECTIFYING CIRCUITS

arrangement it is possible to use a generator having one terminal earthed. This is the Cockcroft-Walton circuit.

Cockcroft-Walton doubler (Figure 6.25)—When A is negative with respect to earth, current flows from the generator, through diode 1, to charge C_1 to a voltage V with the polarity shown.

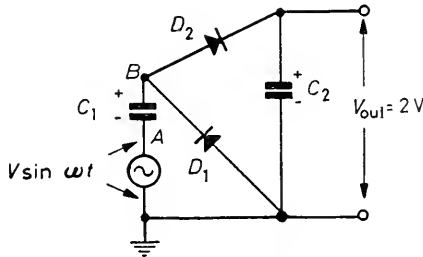


Figure 6.25

When A moves positive with respect to earth, diode 1 is cut off, and as A reaches its peak voltage the potential of B above earth is equal to: generator voltage + C_1 voltage = $2V$; diode 2 conducts and allows C_2 to charge to this potential difference. Notice that the potential at B fluctuates between zero, when A is at peak negative, to $+2V$, when A is at peak positive.

Cockcroft-Walton quadrupler—Suppose we add a further section, comprising two capacitances and two diodes to the top of the doubler circuit already described, as in Figure 6.26. When B swings down to earth potential, diode 3 conducts, allowing C_3 to charge from C_2 to a voltage $2V$. B now swings up to $+2V$, cutting off D_3 and carrying the potential at F up to $+4V$. D_4 then conducts, charging C_4 so that the voltage across C_4 and C_2 in series is also $4V$.

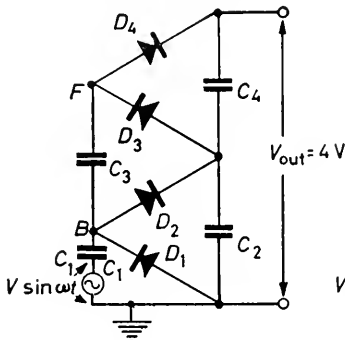


Figure 6.26

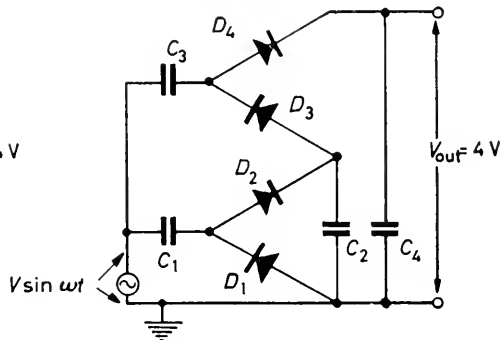


Figure 6.27

It is possible to multiply the generator voltage by larger factors by stacking up further capacitances and diodes in a similar manner. However, due to losses in the diodes the law of diminishing returns operates strongly at high orders of voltage multiplication and the proportional discrepancy between the desired output voltage and the output voltage actually achieved becomes serious. Sextuplers are uncommon.

The Cockcroft-Walton quadrupler is sometimes seen in an alternative form, shown in Figure 6.27. The capacitances can be smaller for a given regulation and ripple, but have to be rated to work at higher voltages.

DIODE CIRCUITS

It sometimes happens that high direct voltages are derived from generators whose output is pulsatory and uni-directional rather than sinusoidal and alternating, and in this case voltage multipliers take a rather different form. There is no difficulty in seeing, in *Figure 6.28*, that if the generator output

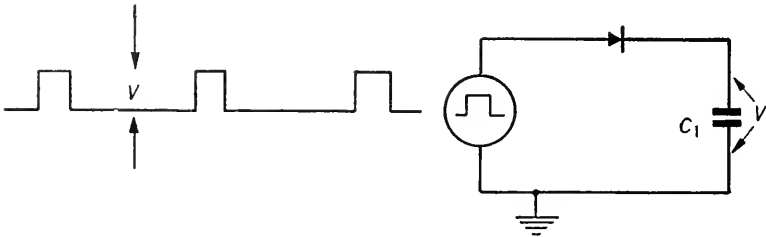


Figure 6.28

rises periodically to V and falls to zero again, then the capacitance will charge to a voltage V also. Now, in *Figure 6.29*, add C_2 and a high resistance. Then between pulses C_2 will charge from C_1 so that the potential at B eventually reaches V also. When the next pulse arrives, A rises to V , carrying B

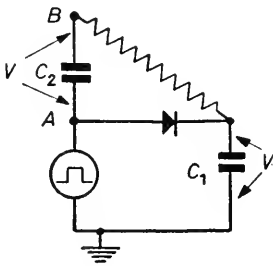


Figure 6.29

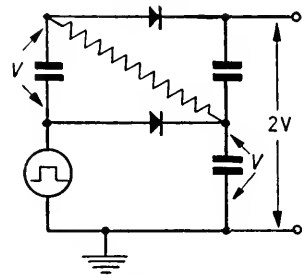


Figure 6.30

up to $2V$, and if we provide another diode and a third capacitance we can arrange for this capacitance to charge from C_2 via the diode (*Figure 6.30*) so that a smooth voltage $2V$ is available with respect to earth. This is the

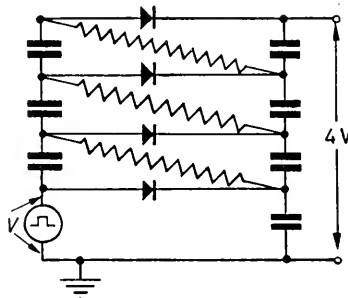


Figure 6.31

voltage multiplier for pulses and as in the Cockcroft-Walton circuit extra voltage may be obtained by stacking further components on top. Thus a pulse quadrupler is shown in *Figure 6.31*.

CATCHING

CATCHING

'Catching' is a diode technique for ensuring that the potential of a point in a circuit shall not exceed a certain prescribed limit, and might be regarded as analogous to buffer stops at the end of a railway siding. Catching circuits are most practicable when the load resistance is high, e.g. a valve, and we shall assume this is the case.

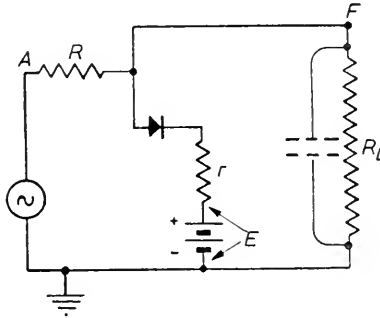


Figure 6.32

In Figure 6.32 we have some kind of generator feeding some kind of signal to a load R_L whose resistance is great. If it is required that the potential at F shall not exceed $+E$ volts, then we arrange a series resistance R , a diode of forward resistance R_f and an E volt battery—or some other direct voltage source—(of internal resistance r) as shown.

When the potential at A is less than $+E$ the diode is cut off and the circuit behaves as if it were not there. Since R is negligible compared with R_L the generator is effectively feeding straight into the load. When the potential at A is greater than E the diode conducts and a potential divider is formed having in the top limb R and in the lower r in series with R_f (Figure 6.33) and further positive excursions of A will be attenuated by the factor

$$\frac{R + r + R_f}{r + R_f}$$

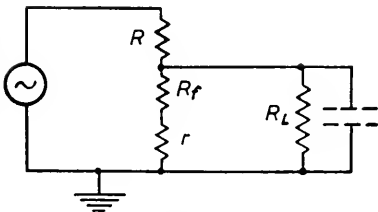


Figure 6.33

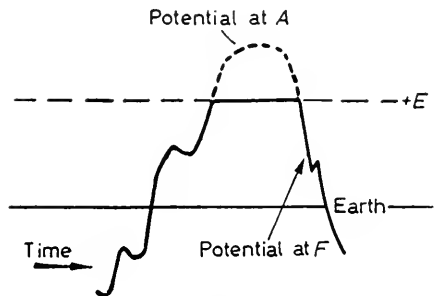


Figure 6.34

Thus in Figure 6.34 the potential at A and the potential at F are seen to part company as soon as the former exceeds $+E$ volts, the latter being prevented from rising far above the 'catching level' $+E$.

DIODE CIRCUITS

The effectiveness of this circuit depends on the attenuation being large when the diode conducts, i.e. R large and $(r + R_f)$ small. It thus appears that any diode, however indifferent its forward resistance, may give as good a performance as is wanted merely by making R large enough. In fact, two considerations impose a restriction on R :

(1) Stray capacitances across the load. These will form a low-pass filter with R , having a cut-off frequency within the required spectrum if R is too large, and frequency distortion will ensue.

(2) If the diode is of the semi-conductor variety the back resistance R_b will form a potential divider with R and there will be a loss of voltage delivered to the load even when the potential at A is within the prescribed limit. Since R_b is a non-linear function of diode voltage there will in addition be introduced amplitude distortion, of an amount proportional to R .

Design procedure is therefore as follows. When the diode is of the semi-conductor kind, choose R at about the geometric mean of R_f and R_b , bearing in mind that in any particular case either efficient catching or absence of amplitude distortion may be more important, then check that loss of high frequencies will not be excessive. If it seems that this will be the case, R will have to be made smaller. If the diode is thermionic, use the highest value of R that will not cause intolerable high frequency loss. If the catching level is to be made adjustable, a potentiometer is added to the circuit (*Figure 6.35*). With this arrangement there is one important point to watch: Efficient catching depended on r , the internal resistance of the battery being low. With the new arrangement we have to substitute $(R_1 R_2)/(R_1 + R_2)$ the effective internal resistance of the source of variable potential. If this is always to be low, whatever the potential setting, $R_1 + R_2$ must also be low, and much power from the battery is wasted in flowing down the potentiometer.

If the output of the generator is a complex of frequencies extending down to zero then this fact has to be accepted and tolerated. However, if there is a lowest frequency ω_L , as is often the case, then we can employ a capacitor (*Figure 6.36*) and the potentiometer can be of as high a resistance as we wish, for the 'catching attenuation' cannot now be less than

$$\frac{\frac{1}{j\omega_L C} + R_f + R}{\frac{1}{j\omega_L C} + R_f} = 1 + \frac{R}{R_f + \frac{1}{j\omega_L C}}$$

In practical circuits the j term can often be made negligible by using quite moderate values of C .

The circuit in *Figure 6.36* might be described as a 'positive catcher', in which case the circuit in *Figure 6.37* is a 'negative catcher', i.e. the potential at F is prevented from going more than a certain amount negative.

CLIPPING

When a positive and negative catcher are combined the result is called a clipper (*Figures 6.38 and 6.39*). The design of these is a perfectly straightforward extension of the procedure for catchers, and requires no further comment.

CLIPPING

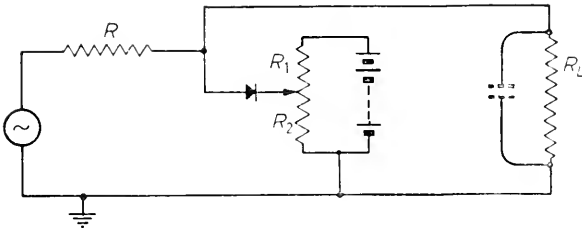


Figure 6.35

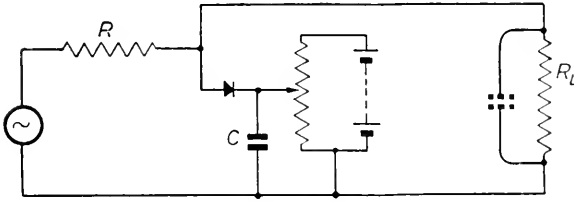


Figure 6.36

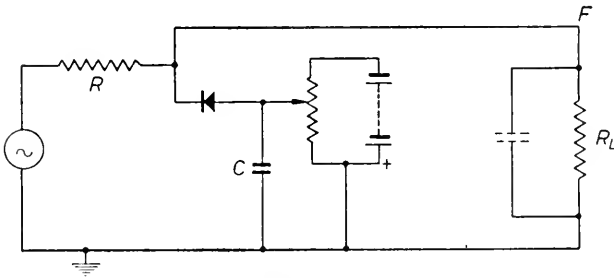


Figure 6.37

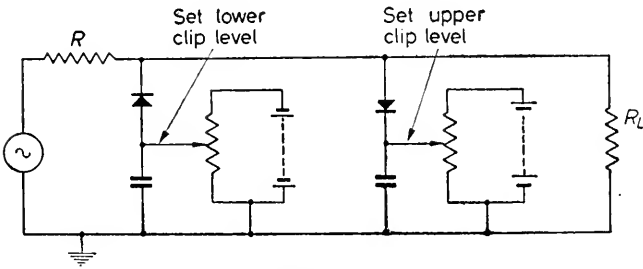


Figure 6.38

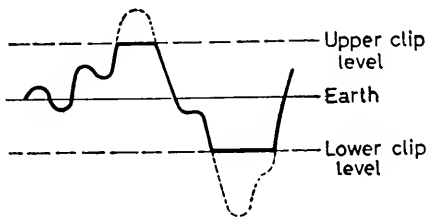


Figure 6.39

DIODE CIRCUITS

CLAMPING

This is a form of electronic switch for suppressing a signal electronically for the duration of a pulse generated elsewhere. As for the catchers and clippers we have a generator feeding a load resistance, which is large, via R . The diodes are connected to a pulse transformer with a centre-tapped secondary, whose primary is fed from the switching pulse generator (*Figure 6.40*).

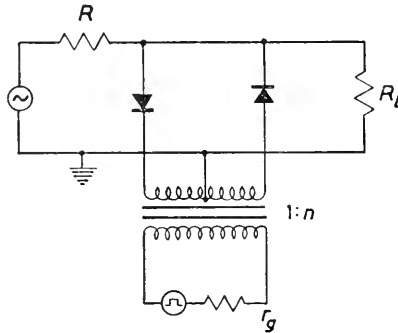


Figure 6.40

In *Figure 6.41a* we have the pulse generator output and in *b* the pulse transformer output. Between pulses the polarity is such as to bias the diodes in the high resistance direction and the signal passes through to the load.

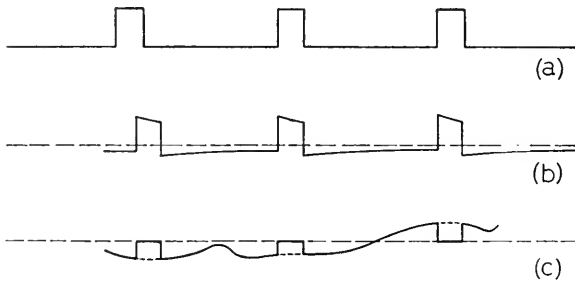


Figure 6.41

On the arrival of the pulses the diode bias reverses, changing them to low resistance, so that the signal output is attenuated as shown in *Figure 6.41c*. The attenuation produced during the clamp period with an ideal pulse transformer is

$$1 + \frac{R}{\frac{R_f}{2} + \frac{r_g}{4n^2}}$$

D.C. RESTORING

Suppose we have a generator feeding a high-pass filter of the simplest kind as shown in *Figure 6.42*, and suppose the waveform of the generator is a train of 'saw-teeth'. This train might have the form of *Figure 6.43a*, where the output moves between $+\frac{1}{2}$ and $-\frac{1}{2} V$, or b , where it is between 0 and $+V$,

D.C. RESTORING

or c , 0 and $-V$. The output of the filter is of necessity of the form of *Figure 6.43a* because this comprises the fundamental saw-tooth frequency and its

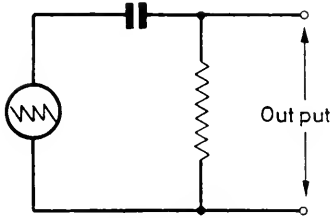


Figure 6.42

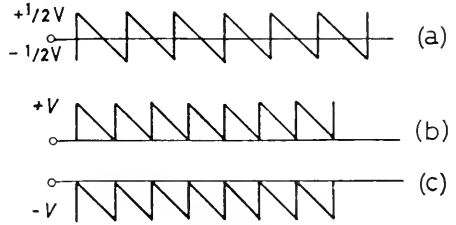


Figure 6.43

harmonics (all of which can be passed without distortion by the filter if RC is large enough) but, unlike *b* and *c*, no direct voltage component. This is to say, the *direct component* present in inputs *b* and *c* is lost because the filter cannot pass zero frequency.

Electronic amplifiers often contain filters of this kind, not because they are wanted as such but because the amplifiers are much easier to design if they are there. The question then arises as to whether the lost direct component can ever be recovered, and the answer is that in general it cannot; but it can be in the two special cases where the positive or negative peaks of the generator output—be they saw-teeth or any other waveform—are at zero potential. That is to say, direct component restoration is possible for the waveforms of *Figure 6.43b* or *c* but not for the waveform in *Figure 6.44*.

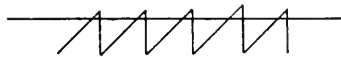


Figure 6.44

The technique of restoration is extraordinarily simple: merely connect a diode across the output as in *Figure 6.45a* to recover inputs of the *Figure 6.43b* type, or as in *Figure 6.45b* to recover the *Figure 6.43c* type. The generators here are supposed to have lost the d.c. component, i.e. these outputs swing between $+V/2$ and $-V/2$. The method of working is merely this: when the generator output in *Figure 6.45b*, say, goes positive, the out-

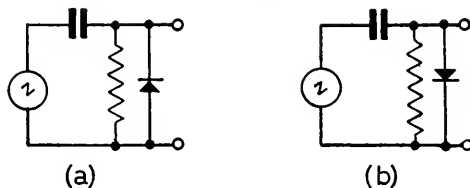


Figure 6.45

put does not also go positive with respect to earth because the diode switches to low resistance and substantially short-circuits the output: diode current flows and quickly charges the capacitance instead. When the generator output goes negative the diode cuts off and the output is carried negative

DIODE CIRCUITS

with respect to earth by an amount $-V$. When the circuit is in equilibrium there is a steady potential difference $V/2$ across the capacitance and diode current flows only at the very positive peaks of the signal, just enough to replace the charge lost by current flowing back round the circuit consisting of resistance and generator.

SHUNT DIODE SIGNAL RECTIFIER

The shunt diode signal rectifier circuit (*Figure 6.46*) has a rather peculiar performance which might lead one to wonder why anyone should want to use it. Its output is a rectified version of the input modulated carrier wave plus the

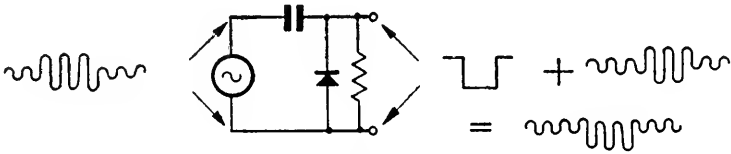


Figure 6.46

modulated carrier wave itself, and usually has to undergo further filtration to remove this carrier component. The reason for its appearance is as follows: many pentode and triode valves include—for a particular purpose in radio receivers—one or two thermionic diodes. These diodes share the same cathode as the main valve and are at a fixed potential. They therefore cannot be used for signal rectification of the series diode type—where both electrodes are ‘live’—but are suitable for shunt rectification where one side of the diode is earthed.

PHASE-SENSITIVE DETECTORS

The function of these valuable devices is perhaps best illustrated by an example:

In recent work F. W. Campbell and W. A. H. Rushton (*J. Physiol.* 130 (1955) 131) have studied the absorption of light by the retina of the eye. To do this, a red and a green beam of light were directed into the eye alternately. The emergent reflected light, which has twice passed through the visual purple, was collected by a photocell. If the amount of visual purple changed, the brightness of the green beam altered with respect to the red. By adjusting a purple wedge (in the path of the green beam) these intensities could be matched (*Figure 6.47a*).

The photocell output is a square wave, and matching is achieved when the amplitude is zero. This is a good method since the characteristics of neither light nor photocell, photocell amplifier or indicating galvo need be well known. Only the wedge must be accurately calibrated.

In order to deflect the galvanometer the signal from the amplifier requires rectification. If it were applied to an ordinary signal rectifier the galvanometer deflection would certainly be proportional to the square-wave amplitude, but would always be in the same direction; there would be no indication as to which way to move the wedge in order to achieve balance. That is, information about the *phase* of the ‘error’ square wave has been lost. What

PHASE-SENSITIVE DETECTORS

is wanted is a signal rectifier which will give a direct voltage output, whose amplitude is proportional to the error wave amplitude, and whose polarity depends on whether the retina beam or the wedge beam is greater. That is,

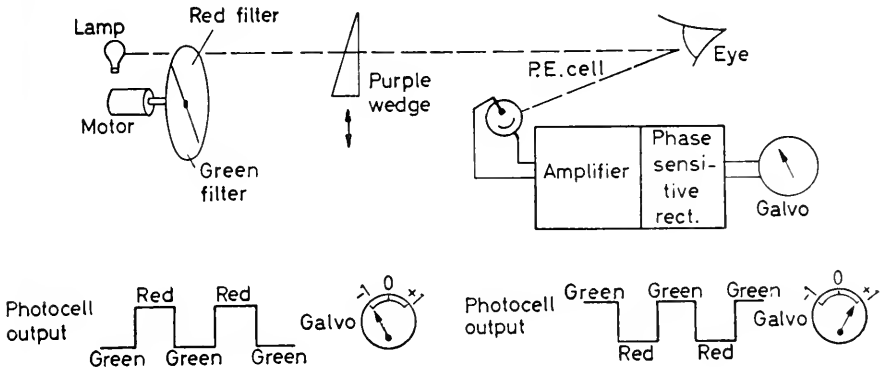


Figure 6.47

a rectifier which identifies which segments of the square wave belong to which beam (Figure 6.47b). This is the purpose of the 'phase sensitive-rectifier'.

To achieve phase-sensitive rectification it is necessary to have an electrical reference wave, usually of square form, which is generated synchronously with the switching of optical beams (Figure 6.48). This may be generated

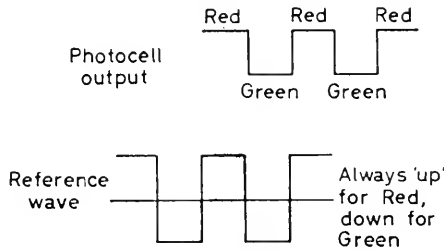


Figure 6.48

by a commutator on the shaft carrying the rotating shutter, or by an auxiliary lamp and photocell system, the light from the auxiliary lamp being interrupted by the same shutter as is used to control the main beams. However it is generated the wave may have to undergo amplification, as an amplitude of the order of 10 V is required for operation of the phase-sensitive rectifier.

The simplest phase-sensitive rectifier is shown in Figure 6.49. The reference wave is applied through a transformer to two similar diodes in series, via similar current limiting resistors. In this circuit and in those to come the diodes are used as switches controlled by the reference wave: when the latter drives current through the diode in the forward direction the diode resistance is low. When the reference wave attempts to drive current through the diode in the reverse direction, the diode resistance becomes high.

In this case it is important to see that, if no input is connected, by symmetry, whatever the potential difference across the transformer happens to be A

DIODE CIRCUITS

and B are at the same potential. Since B is earthed, A must be at earth potential too. The action of the reference wave on the diodes is alternately to make high and low the resistance of the path from A to earth. Now let

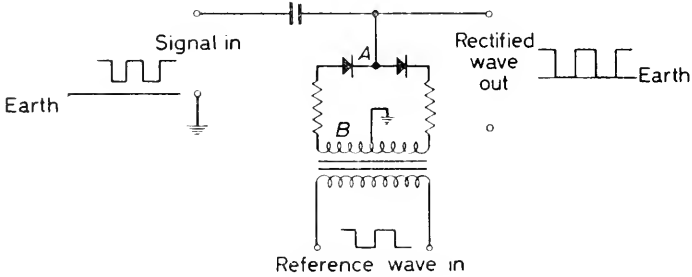


Figure 6.49

the signal square wave be applied to the left-hand plate of the coupling capacitor. If the capacitor is large, the same voltage waveform must be present on the right-hand plate too, but as either the negative or positive sectors are earthed by the diodes—depending on the phase of the input—

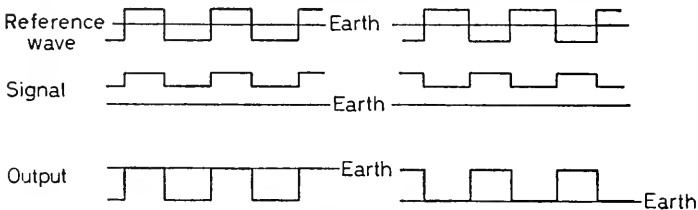


Figure 6.50

the output wave, as it were, ‘stands’ or ‘hangs’ from earth potential (Figure 6.50). By passing the output through a low-pass filter the average value may be extracted, and this is the output required.

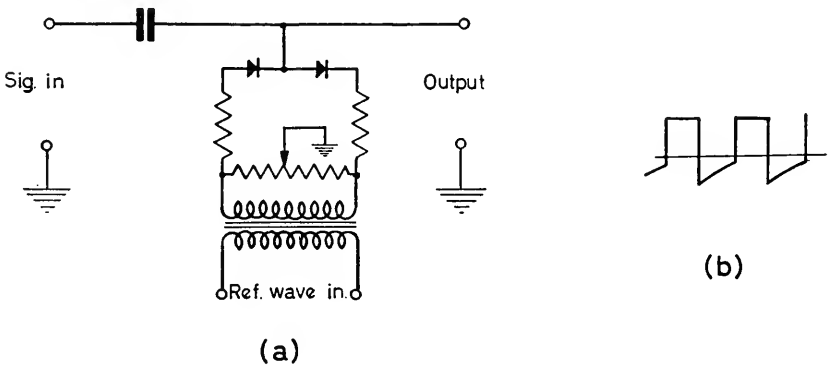


Figure 6.51

This simple circuit has two snags: (1) if the diodes are not perfectly similar the output is complicated by a component of the reference wave; (2) the path to earth is not of very low resistance, since it involves the current limiting resistors. The former objection can be overcome by replacing the

PHASE-SENSITIVE DETECTORS

centre-tap by a potentiometer (*Figure 6.51a*), which is then adjusted until, with no input applied, the reference wave produces no spurious signal at the output. With this arrangement the resistance to earth is still not very low, and the effect of this is to cause the output wave form not to lie wholly on one or other side of earth potential (*Figure 6.51b*). Upon taking the average, the output is in consequence reduced.

A better scheme is to use four diodes connected so that all are switched to low resistance together (*Figure 6.52*): this device is known as the Cowan

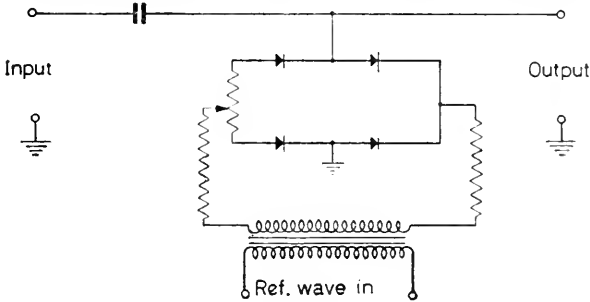


Figure 6.52

bridge. A potentiometer is still necessary to eliminate the reference wave from the output, but by putting it in the left-hand side of the bridge, as shown, there is a low resistance path to earth straight through the right-hand diodes, so the rectification is efficient.

What might be called a 'full-wave' phase-sensitive rectifier is shown in *Figure 6.53* and this is the form used by Rushton and Campbell. Two trans-

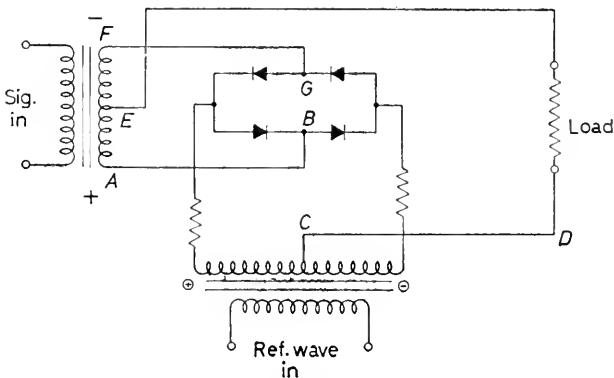


Figure 6.53

formers provided with centre-tapped secondaries are required. The action of the reference wave is to make first the upper pair then the lower pair of diodes conducting. Thus, with the polarities shown, the upper diodes are cut off and no current leaves or enters at G. The lower diodes conduct and current flows along the path *ABCDE*, i.e. up the load. The next half-cycle, polarities reverse and the current path is *FGCDE*, i.e. still up the load, and the action is seen to be full-wave. The load current clearly reverses only if the phase relationship of reference wave and signal wave reverses.

SOFT VALVES

'Soft' valves are those in which the electrodes are sealed in a glass envelope containing either an inert gas or mercury vapour at a reduced pressure. The list is nearly but not quite complete if we tabulate them in the following manner:

	<i>Hot Cathode</i>	<i>Cold Cathode</i>
Two-electrode	High power rectifier	Voltage reference tube Stabilizer tube Difference diode
Three-electrode	Thyratron	Trigger valve Primed stabilizer

Electrical conduction in metals is possible because of the presence within the matrix of the material of a large number of free, unbound electrons. Due to the thermal agitation of the molecules of which the metal is composed electrons are continually being shot out of the metal altogether (thermionic emission); usually they merely fall back into the metal again. The average rate at which electrons are emitted is not unexpectedly a function of the temperature of the metal and of a constant for the metal itself. This is of the form $I = AT^2e^{-B/T}$ where T is the absolute temperature, a relationship known variously as Richardson's or Dushman's equation.

COLD CATHODE DIODE

If two electrodes are connected to a source of voltage and immersed in an inert gas, then an electron casually emitted at the cathode experiences an attraction towards the anode and will not necessarily fall back at once into the cathode again: instead it may hit a gas molecule. If the electron hits the gas molecule sufficiently hard it will ionize it, producing another electron and a positive ion. This molecular disruption will occur when the potential difference across the device exceeds a certain critical value called the ionization potential. Positive ions so created set off towards the cathode and electrons towards the anode. A minute current flows through the valve. If the potential difference is further increased a point is reached at which the heavy positive ion on hitting the cathode knocks out further electrons and the chain reaction is established at which the cathode and anode become joined by a glowing, conducting column. The potential difference at which this occurs is the 'striking voltage'.

A chain reaction is an unstable state of affairs and requires 'moderating'; if a cold cathode diode or any other soft valve be connected to a constant-voltage supply above the striking potential the current rises at once to an indefinitely high value, at which the device is destroyed. To remedy this a

COLD CATHODE DIODE

series stabilizing, or 'ballast' resistance must always be used (*Figure 7.1*). It is then found that, on applying a voltage larger than the 'striking' voltage, the glow is established and the potential difference across the diode falls to a lower 'running' voltage. The arrangement is now stable because any tendency

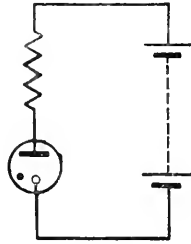


Figure 7.1

for the glow to thicken and the current to increase is met by an increased voltage drop in the ballast and hence a reduced difference of potential applied to the valve. The difference between the striking and running voltage is called the 'voltage differential' of the tube. Tubes having a large differential will be called 'refractory'. With proper care in manufacture the running voltage can be made very constant despite considerable changes in the glow current. This is because as the current is varied the thickness of the glow column varies, but the potentials involved are unchanged. This property is made use of in the voltage reference and voltage stabilizer tubes. The upper limit of current which may be passed through these tubes is set by the point at which the glow column completely surrounds the cathode surface. No further increase in current is then permissible or the potential difference across the tube will suddenly rise. The cathode is then liable to destruction by ionic bombardment.

Voltage reference tube

This is for producing a stable reference voltage from a supply which is subject to fluctuations (*Figure 7.2*). It is not intended that power be drawn

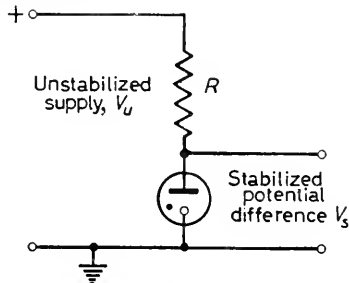


Figure 7.2

from the output terminals (though small amounts may be); the device is intended to be used in the same sort of way as a Weston standard cell. Reference tubes are designed to operate with low glow currents of the order of a milliamp.

The running voltages of reference tubes lie within the range 60–100 V

according to type. Higher voltages may be had by connecting tubes in series (Figure 7.3). They may be of different types if this helps to get the required running voltage, but design is easier if they are all intended by the makers to work over a similar current range. Having chosen the tube or tubes to use attention must be given to the minimum value V_u can have in order that the tube or tubes shall strike. For a single tube this is merely the quoted striking voltage. For several tubes, proceed as follows. Arrange the

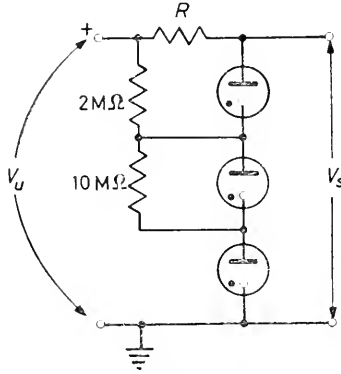


Figure 7.3

tubes in order so that those with the greatest difference between their running and striking voltages are at the bottom and those with the least are at the top. Now take a set of high resistances, in number one less than the number of tubes, the lowest of the order of megohms and of values such that each is, say, 5 times the value of its predecessor. 2 megohms and 10 megohms would be suitable for a chain of 3 tubes. Connect them as in Figure 7.3. Now consider what happens when V_u is connected. Let the running voltages of the tubes be V_1 , V_2 and V_3 . At first the whole of V_u is applied to the bottom, most refractory tube by way of the high resistances, ensuring that it strikes. Then we have $V_u - V_3$ applied to the upper two tubes, but because of the potential divider now made by the resistances, $\frac{5}{6}$ of it is applied to V_2 , ensuring that it strikes. Finally we have $V_u - (V_3 + V_2)$ applied to the upper, least refractory tube, and this will strike if $V_u - (V_3 + V_2)$ is greater than its striking voltage, or if $V_u - (V_1 + V_2 + V_3)$ exceeds the voltage differential of this tube. The rule is that the tubes will strike if the difference between the unstabilized input and the stabilized output voltages exceeds the difference between the running and striking voltages of the least refractory tube, provided the tubes are arranged to strike in the right order.

Having found $V_{u\text{min}}$, we can look about for some point in the power pack from which to tap off V_u . Stabilizers are not perfect and in fact there is some change of voltage across them when the glow current changes, so that as far as *increments* of voltage and current are concerned they may be allotted an equivalent resistance. Thus if a reference tube voltage changes from 69.9 to 70.1 volts for a glow current change from $\frac{1}{2}$ to $1\frac{1}{2}$ milliamps, its incremental resistance is $0.2/10^{-3} = 200$ ohms. Note that this is a non-linear device and the incremental resistance is not the same as the absolute

COLD CATHODE DIODE

resistance, which is $70 \text{ V}/1 \text{ mA} = 70 \text{ k}\Omega$. Referring to *Figure 7.4*, it is clear that the goodness of the stabilization depends on the attenuation of increments in V_u , that is, on R being much larger than 200 ohms.

There is, however, a catch here. Suppose we choose a tube current of 1 mA as lying nicely in the centre of the makers' recommended range, and suppose $V_s = 100 \text{ V}$. A 250-fold reduction in fluctuations of V_u sounds attractive, so let us aim for this. Then R must be 50,000 ohms and the

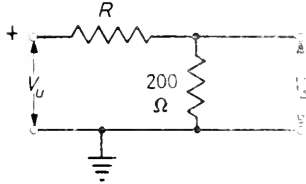


Figure 7.4

voltage drop across it will be $50,000 \text{ ohms} \times 1 \text{ mA} = 50 \text{ V}$, so $V_u = 150 \text{ V}$. Now let us try to do four times better. For a 1,000-fold reduction we want R to be 200,000 ohms, hence the drop is 200 V and so $V_u = 300 \text{ V}$. Suppose now V_u undergoes a 10 per cent upward change in value. In the first case it alters from 150 to 165 V, so the increment is 15 V and the proportion that is added to V_s is $1/250 \times 15 = 0.06 \text{ V}$. In the second case the change in V_u is 30 V and the proportion of the change added to V_s is $1/1,000 \times 30 = 0.03 \text{ V}$, i.e. only twice as good. If we aim for a 4,000-fold reduction we get $R = 200 \text{ ohms} \times 4,000 = 800,000 \text{ ohms}$, and the drop across it would be 800 V and V_u would have to be 900 V. A 10 per cent increase would then be a change of 90 V, and the fraction appearing at the output is $1/4,000 \times 90 = 0.0225 \text{ V}$, an improvement hardly commensurate with requiring 900 V to produce 100.

In point of fact there is little to be gained by making V_u more than two or three times V_s , and the proper thing to do is to get V_u from some convenient point in the circuit and *then* to decide on the glow current, rather than the other way round. The performance of the circuit is improved if R is increased, and R can be increased until the tube current approaches the lower limit advised by the makers, below which the tube performance will become unsatisfactory. It is a good general rule in any case with soft valves to have low currents if possible, because the life falls off inversely as a high power of the current—about the fifth.

Voltage stabilizer tubes

As with the voltage reference tube, one or more of these are arranged in series to produce the required steady output voltage V_s from a supply voltage V_u subject to fluctuation, but in this case it is intended that they be used to stabilize the voltage supplied to some sort of load, represented by R_L in *Figure 7.5*. If the load resistance is also liable to fluctuations, such that the load current varies by an amount δI_L , and if when the load is maximal a 'keep alive' current I_{\min} flows through the tubes, then clearly when the load is minimal the tube current must rise to $I_{\min} + I_L$. It emerges that voltage stabilizer tubes may have to pass appreciable currents, typically 40–100

SOFT VALVES

milliamps. Whilst they are therefore similar in principle to reference tubes, they are much larger and more robust affairs.

To ensure that the tubes will strike we arrange them in order, with high resistances across the less refractory ones, and apply the same rule as for reference tubes, except that the voltage seen by the tubes is now not V_u but $V_u\{R_L/(R + R_L)\}$.

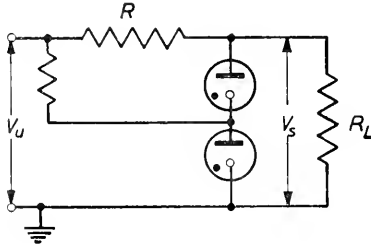


Figure 7.5

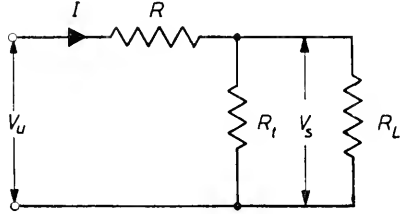


Figure 7.6

If the string of stabilizer tubes has an incremental resistance R_t , then so far as the reduction of the effect of fluctuations is concerned the equivalent circuit is *Figure 7.6*. We have now to consider two aspects of stabilizer performance. ‘Forward’ stabilization is the tendency of the arrangement to maintain V_s despite fluctuations in V_u . ‘Backward’ stabilization is the tendency of the arrangement to maintain V_s despite fluctuations in R_L , that is, despite variations in load current. Clearly the forward stabilization is good if R_t or R_L are much smaller than R . Also the backward stabilization is good if the impedance looking back towards the stabilizer from the load is low compared with the load resistance. Therefore for good all-round stabilization we need R and R_L large and R_t small. On putting in some figures it turns out that the design of stabilizer circuits is surprisingly complicated and may even be impossible. For, calling the striking potential of the tube or string of tubes V_i , we have: that they may just strike, $V_i = V_u(R_L)/(R_L + R)$. Also that once they have struck $V_u = IR + V_s$. Eliminating R between these equations

$$V_i \left(\frac{V_u - V_s}{I} + R_L \right) = V_u R_L$$

or

$$V_u = V_i \frac{IR_L - V_s}{IR_L - V_i}$$

Looking at the denominator of this it is evident that if the striking voltage V_i is so high that it approaches IR_L then the expression goes to infinity, which means that the tubes will not strike however high we make V_u . The arrangement is probably intolerably inefficient if V_u is greater than twice V_i , that is, one should aim for

$$2 > \frac{IR_L - V_s}{IR_L - V_i}$$

or

$$IR_L > 2V_i - V_s$$

COLD CATHODE DIODE

Suppose we wish to design a stabilizer to supply 210 V at 0–70 mA. Then the load resistance which corresponds to maximum current is $210 \text{ V}/70 \text{ mA} = 3,000 \text{ ohms}$. Suppose a string of stabilizers have been chosen which will strike at 255 V. Then $2V_i - V_s = 510 - 210 = 300 \text{ V}$, and IR_L must be greater than this. Since $R_L = 3,000 \text{ ohms}$, I must be 100 mA or greater. Let it be 10 per cent over the critical value, i.e. 110 mA, for safety. Taking

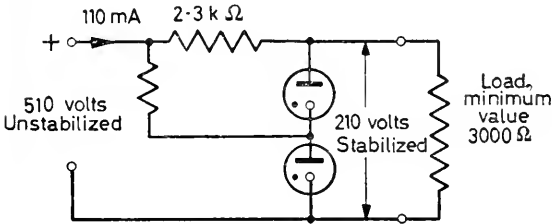


Figure 7.7

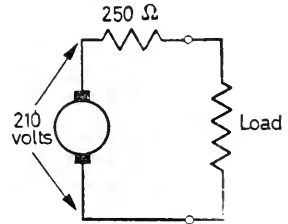


Figure 7.8

$V_u = 2V_i = 510 \text{ V}$, the provisional design is in *Figure 7.7*, and the ballast resistance is $110 \text{ mA} \times 210 \text{ V} = 2,300 \text{ ohms}$. At full load 40 mA is delivered to the tubes and 70 mA to the load. On no load all 110 mA pass through the tubes which must be rated to pass this current. The practicability of the design depends on whether tubes can be found to pass up to 110 mA. If they can only be found to pass up to, say, 85 mA then the current output is restricted to 25–70 mA. Let us see how it behaves. On first applying V_u , the potential difference across the load rises to $3,000/(3,000 + 2,300) \times 510 = 278 \text{ V}$, so that the tubes strike with 23 V to spare. Thereafter, for an incremental tube resistance of 250 ohms—a typical value—the forward stabilization ratio $\delta V_s/\delta V_u$ is, exactly

$$\frac{250 \cdot 3,000}{250 + 3,000} \\ 2,300 + \frac{250 \cdot 3,000}{250 + 3,000}$$

which is not far from $250/(2,300 + 250)$ which is approximately equal to $1/10$.

Such a performance is greatly inferior to that of a voltage reference tube as regards stability of the output, but is often worth having. The backward stabilization will be poor, for the equivalent circuit of the arrangement is substantially that of *Figure 7.8*, and there will be a change of output voltage of $70 \text{ mA} \times 250 \text{ ohms} = 17.5 \text{ V}$ between 0 and 100 per cent of full load. The tube life is likely to be short. Tubes cannot efficiently be run in parallel to help matters.

Now consider a stabilizer to supply 210 V at 0–7 mA, using the same tubes. The full-load load resistance is now 30,000 ohms. As before $IR_L = 300 \text{ V}$, so I must be 10 mA or more. If it is 10 mA and the load current is 7 mA, only 3 mA would flow through the tubes and it is doubtful if they would work properly at such a low current. Let us assume the lower current limit for these hypothetical tubes is 10 mA. Then I would have to be 17 mA

SOFT VALVES

and the ballast resistance is $210 \text{ V}/17 \text{ mA} = 12,350 \text{ ohms}$ for a V_u of 510 V. The forward stabilization ratio will be

$$\frac{250}{250 + 12,350} = \frac{1}{50}$$

and the variation in output voltage between 0 and 100 per cent full load will be $7 \text{ mA} \times 250 \text{ ohms} = 1.75 \text{ V}$. Clearly both backward and forward stabilization are much improved by reducing the maximum current demands on the device.

Difference diode

In the design of voltage reference and voltage stabilizer tubes every effort is made to make the striking voltage as little above the running voltage as possible, as this facilitates the design of the ancillary circuit, as we have seen. In the difference diode, on the other hand, the striking voltage and running voltage are deliberately made widely different. This property can be made use of in various ways, of which two will be mentioned.

Bi-stable circuit—A bi-stable circuit is one which has two distinct stable conditions and may be altered from one to the other by appropriate electrical pulses. The difference diode bi-stable circuit is extremely simple and appears in *Figure 7.9a* and *b*. Its two states are simply 'aglow' and 'not aglow'.

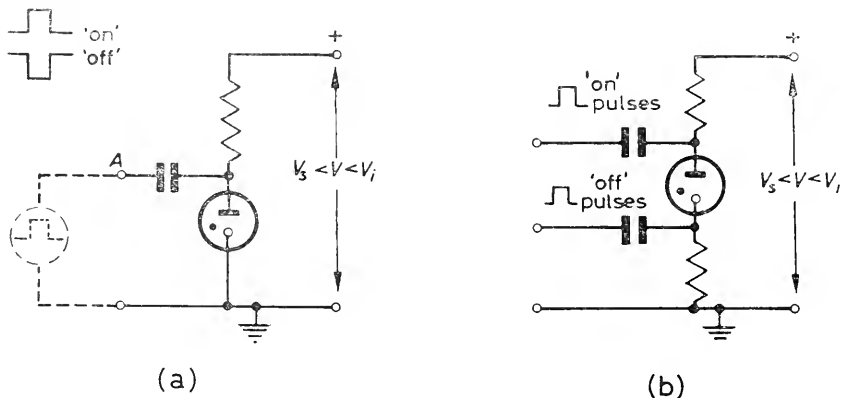


Figure 7.9

In *Figure 7.9a*, we have a difference diode and ballast resistance fed from a direct supply voltage which is greater than the running voltage but less than the striking voltage of the tube. Control pulses are introduced (via a capacitance) with respect to earth by a pulse generator of some kind, connected at *A*, as shown.

Suppose the tube be at first extinguished. Then if a positive pulse be applied to the anode sufficient to carry it above the striking potential, the glow is initiated and the anode falls to the running potential and the circuit remains in this 'aglow' condition indefinitely. However if now a negative pulse is fed in at *A*, the anode is carried below the running voltage and the glow is extinguished. On the cessation of the pulses the anode potential returns to $+V$ and the circuit is once more stable in the 'not aglow' condition until a further positive pulse arrives at *A*.

COLD CATHODE DIODE

Another version of the circuit is in *Figure 7.9b*, for switching pulses of the same polarity. The glow is switched on by making the anode temporarily extra positive, increasing the anode-cathode potential difference as before, but is switched off by making the cathode temporarily extra positive, thus reducing the anode-cathode potential. Alternatively, of course, the circuit may be switched 'on' by negative pulses at the cathode and 'off' by negative pulses at the anode.

Relaxation oscillator—Relaxation oscillators are those whose output waveform is non-sinusoidal. An extremely simple relaxation oscillator can be made with a difference diode, a capacitance, a resistance and a direct voltage supply, (*Figure 7.10*), which exceeds the striking potential of the diode.

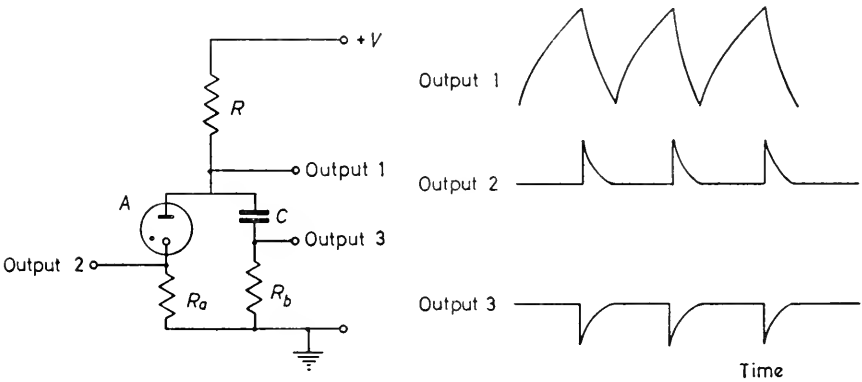


Figure 7.10

On connecting V , the capacitance charges via resistances R and R_b and the potential at A —Output 1—rises exponentially towards $+V$. The tube is non-conducting, so no current flows through R_a and there is no voltage drop across it, so the P.D. between A and earth also appears across the diode. When it reaches the striking voltage of the diode the latter becomes suddenly a conductor of low resistance, and the capacitance discharges quickly, passing current anti-clockwise round the circuit comprising the tube, R_a and R_b . As a result of this, the potential at A falls rapidly toward the running voltage of the tube, and eventually falls below it. The discharge in the tube is then extinguished and recharging of the capacitance begins and the cycle is repeated. An output in the form of positive going 'pips' may be taken off across R_a , or negative going pips from across R_b . The sum of the peak amplitudes of the two sets of pips is given by the voltage differential of the tube, and the sizes of the two sets are in the same ratio as R_a to R_b . If only negative going pips are required, R_a may be omitted, and vice versa, but in general R_a and R_b must be chosen so that

$$R_a + R_b > \frac{\text{Striking voltage} - \text{Running voltage}}{\text{Maximum permissible tube current}}$$

otherwise when the capacitance discharges the tube current will be excessive and the cathode liable to destruction.

The output at A has the form of a saw-tooth of peak-to-peak amplitude also equal to the tube voltage differential. If C is small the discharge is

SOFT VALVES

rapidly accomplished. If R is large the charging phase may be made indefinitely slow. If V is made large the charge is made quicker—an effect which may be combated by further increase in R —but, more important, only the



Figure 7.11

first part of the exponential is used and this is approximately linear. The output is thus of the 'quasi-triangular' form of Figure 7.11. Quasi-triangular waveforms occur frequently in electronics.

HOT CATHODE DIODE

In the cold cathode diode conduction only becomes appreciable when the energy possessed by positive ions, created by a few chance collisions, becomes sufficient to knock further electrons out of the cathode to produce a sustained discharge; this only happens when the potential difference is rather large, between 60 and 150 V, according to the type of tube. If the cathode temperature be raised to dull red heat, usually by making it in the form of a ribbon through which a special heating current is passed, there is an adequate supply of electrons thermionically emitted to represent a very considerable current. If the anode be made negative with respect to the cathode, these electrons are all turned back to the cathode and no current flows. If the anode be made slightly positive with respect to the cathode, some electrons pass to the anode, and others cloud around the cathode to form a negative 'space charge' which tends to prevent the emission of further electrons, and a small current flows. This phase in the operation of the valve is known as 'space charge limited'. When the anode potential is raised further to a critical value—the

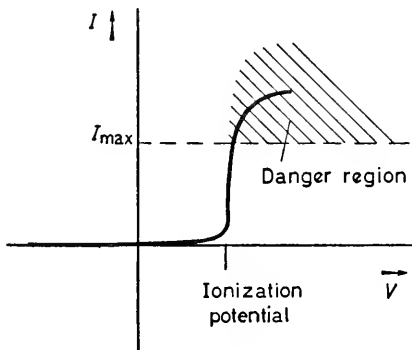


Figure 7.12

ionization potential—which is between 10 and 25 V, depending on the gas used, the average electron energy is suddenly sufficient to ionize the gas filling. When this happens the positive ions travel towards the cathode and neutralize the space charge. There is now nothing to prevent any current being passed by the diode that the external circuit dictates, up to the limiting

COLD CATHODE TRIGGER VALVE

emission as given by the Richardson equation, and over this range the anode-cathode voltage is constant at the ionization potential.

Plotting these results in *Figure 7.12* it is clear that the device is a diode in the sense of Chapter 6. The shaded region requires comment. If any attempt is made to draw more current than the rated maximum from the valve, the extra electrons to represent this current can only be emitted by knocking them out by ions, which means that the potential difference across the valve must rise. Not only does this represent an increase in forward resistance, nearly always a bad thing, but also the cathode will be quickly destroyed by the excessive ionic bombardment. In *Figure 7.13* the characteristics of soft

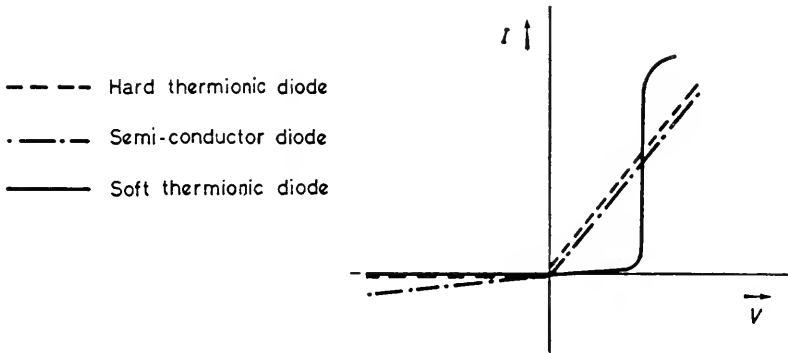


Figure 7.13

thermionic diodes are compared with those other types of approximately the same rating. It is evident that at high currents the soft valve is more satisfactory because the voltage drop across it is lower. For this reason the application of these valves is in power rectification of high currents; they tend therefore to be found in circuits employing the choke-input system. In using them two precautions are important:

(1) The current must be limited to the rated maximum by some protective device such as a fuse or magnetic cut-out.

(2) The cathode heater must be switched on and the cathode given time to warm up before the high voltage supplies are switched on. If this is not done there will be a period at which the cathode temperature is not yet sufficient to supply all the load current thermionically, and the remainder will be derived by ionic bombardment, with consequent damage to the valve.

COLD CATHODE TRIGGER VALVE

The cold cathode trigger valve may be regarded as a difference diode to which a third, trigger, electrode has been added. In *Figure 7.14*, if $+V$ is between the striking and running voltage of the tube, discharge may be initiated by applying a positive pulse to the trigger electrode. The trigger is given a positive 'bias' by a high-resistance potential divider across the power supply such that the tube just does not 'fire': a small pulse only is then sufficient to fire the tube. Once the discharge is established, it cannot be extinguished by the trigger electrode; other means are necessary to achieve this.

In the so-called 'high speed' trigger valve it is necessary that the discharge

SOFT VALVES

current shall build up quickly to its full value, and to facilitate this it is necessary that there be an adequate supply of electrons near the cathode. To achieve this the tube contains a subsidiary cathode and anode between

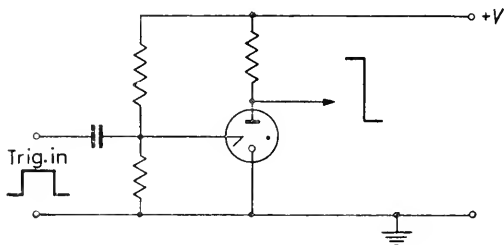


Figure 7.14

which a 'keep alive' discharge is permanently maintained. The light output from the keep-alive discharge ensures a sufficiency of photoelectrons in the neighbourhood of the main cathode.

PRIMED STABILIZER

This is a voltage stabilizer tube to which a 'primer' electrode has been added. The primer is returned to a positive supply potential via a high resistance (*Figure 7.15*). The advantage of the primed stabilizer is a reduced striking

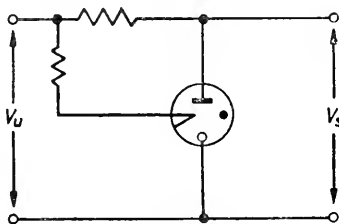


Figure 7.15

voltage for a given running voltage, making the ancillary circuitry easier to design.

THYRATRON

If a wire grid is interposed between the cathode and anode of a soft hot-cathode diode, and the connection to it is brought out, the performance is modified in the following manner.

Suppose the grid be made a few volts negative with respect to the cathode, the electrical field produced tends to reinforce the space charge, and assists it in returning electrons to the cathode and in retarding those which escape from it. The average electron velocity is therefore lower and an anode potential which normally would produce ionization now fails to do so. To produce ionization the anode potential has to be made more positive by an amount proportional to that by which the grid has been made negative. The ratio between negative grid 'bias' and critical anode potential to cause ionization is a constant and is called the 'control ratio' of the thyatron.

When once ionization does occur current flows through the valve,

THYRATRON

determined largely by the limiting resistance R (Figure 7.16), and the voltage across the valve falls to the ordinary ionization potential, about 15 V. The current is in fact $(V - 15)/R$; a typical small thyatron will pass up to 100 milliamps. As a result of the current of positive ions travelling towards the cathode region, not only is the space charge cancelled, but also the effect of the negative grid, and it is impossible to shut off the discharge by making the grid more negative since the only effect this has is to attract more ions into the cathode region. We say that the grid 'loses control'. Instead of fixing

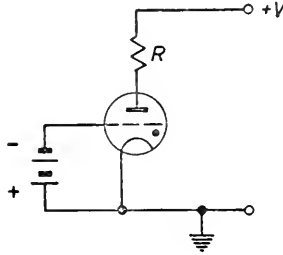


Figure 7.16

the grid bias and raising the anode potential until ionization occurs, we can of course fix the anode potential at some value greater than the ionization potential and initiate conduction by raising the grid potential from very negative to [anode potential]/[control ratio]. When this is done (bearing in mind that it subsequently loses control) the grid is said to be being used as a 'trigger' and—extending the analogy—when ionization occurs the thyatron is said to 'fire'. The only way to stop the discharge again is to reduce V below the ionization potential.

Precautionary measures to be taken with thyratrons are similar to those for soft hot-cathode diodes. The cathode must be allowed to warm up before the anode potential is applied, and the discharge current must be held within the makers' upper limit, for fear of destroying the cathode.

Thyatron relaxation oscillator—One difficulty with the difference diode relaxation oscillator is that the amplitude of the three outputs cannot readily be controlled electronically. Output (1) is fixed by the voltage-differential of the tube, and outputs (2) and (3) by the relative sizes of R_a and R_b (Figure 7.10). The thyatron relaxation oscillator can be controlled by a potential on the thyatron grid (Figure 7.17). The action is similar to the difference diode oscillator, but whilst the capacitance always discharges to the ionization potential, the voltage to which it charges can be made greater by making the grid more negative. Unfortunately the frequency of oscillation is also affected and has to be corrected by an appropriate alteration in R .

In designing these circuits it must be remembered that $R_a + R_b$ must be chosen so that the peak thyatron current is within the safe limit when the grid is as much negative as it is ever likely to be, for it is then that the capacitance charges to the highest voltages and the discharge currents are greatest, that is

$$R_a + R_b \geq \frac{\text{Maximum anode voltage} - \text{Ionization potential}}{\text{Maximum permissible thyatron current}}$$

SOFT VALVES

Thyratron switch—The criterion of a good ‘on-off’ switch is that when the switch is ‘off’ the resistance be extremely high and when the switch is ‘on’ the resistance be low, that is, that the potential difference across it be small

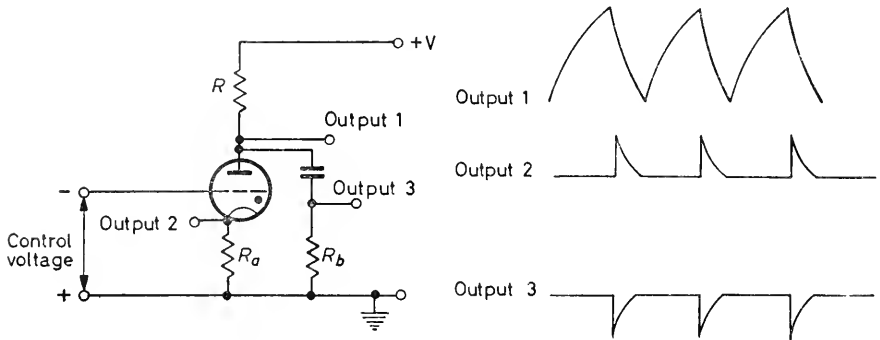


Figure 7.17

when it is passing the load current. The thyratron is a good electronically controlled switch (*θvρα*—a door), for provided it is capable of passing the load current the voltage dropped across it cannot exceed some 15 V, which is not much in apparatus commonly supplied with voltages of around 500.

The usefulness of the thyratron in switching direct current supplies is somewhat reduced by its inability to switch ‘off’. In Figure 7.18 we have a direct voltage generator supplying some kind of load, an electromagnet or a motor or a lamp, perhaps, and the circuit can be made by removing a heavy negative bias from the thyratron grid. However it must be broken in some other way, perhaps manually with a push button. Depending on the application, this may or may not put the thyratron out of court.

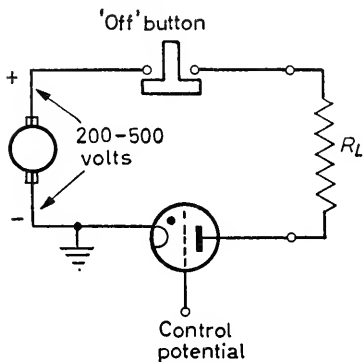


Figure 7.18

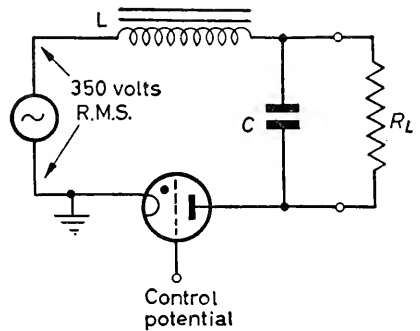


Figure 7.19

If the supply is a.c., then the grid can be used both for switching on and off. For example, in Figure 7.19 the thyratron acts both as a switch and as a rectifier. If the grid is at cathode potential we have virtually a soft diode choke input rectifying circuit feeding d.c. + ripple to the load. When the upper generator terminal is negative the thyratron anode is negative with respect to the cathode and no conduction can occur. When *A* is positive the

THYRATRON

situation is reversed and the thyatron conducts. The conduction period is approximately half a cycle. If the grid is made more negative than

$$\frac{(2)^{1/2} \times 350}{\text{Control ratio}} \text{ V}$$

then the thyatron is prevented from firing at any time, i.e. the switch is 'off'. If the grid is made rather less negative than this, conduction will occur over less than half a cycle and an intermediate amount of power is supplied to the load. This is the theory of the grid-controlled rectifier, which is complicated and probably of little interest in connection with biology. It is worth mentioning if only to draw attention to the fact that it exists, and to emphasize that in the design of simple on-off thyatron switches anomalous results may ensue if care is not taken that the shut-off bias is large enough.

HARD VALVES

Hard valves are those in which the electrodes are sealed in a glass envelope containing a vacuum. The object of this is to exclude as far as possible the complicating effects of ions in the valve. Hard valves are described according to the number of electrodes they contain: diode, triode, tetrode (or screened grid), pentode, hexode, heptode (or pentagrid), octode and nonode. Of these the screened grid, hexode, heptode, octode and nonode valves are of little interest in general electronic work, being used for specialized radio applications; the important hard valves for electrobiology are the diode, triode, beam tetrode and pentode.

Hard valves employ thermionic emission from a heated cathode. The cathode is heated by current supplied from a special auxiliary circuit which for clarity is usually left out of circuit diagrams. The heating may be achieved in either of two ways.

DIRECTLY HEATED CATHODE

The cathode is in the form of a thin wire or ribbon of tungsten which is heated by the passage of a suitable current through it (*Figure 8.1*). It is usually coated with a material which emits electrons richly at relatively low

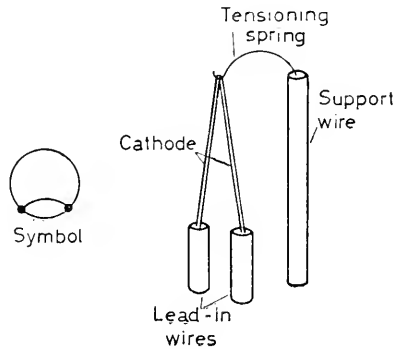


Figure 8.1

temperature—dull red heat—thus reducing the amount of heater power which has to be supplied. This material is usually a mixture of barium oxide and strontium oxide. The thermal capacity of the cathode is low so that to achieve a steady emission the heater current has usually to be direct; a few valves having particularly robust cathodes—capable of holding more heat—are designed for direct heating from a.c. They are mostly power rectifier diodes, for in power rectification there is ripple in the output anyway, and a little more due to 100 cycle fluctuation in cathode temperature does no harm.

OTHER ELECTRODES

The advantages of direct heating are its simplicity and the rapidity with which the cathode reaches its operating temperature. The disadvantages are: (1) that the heater supply must necessarily be at cathode potential, so that a number of valves having their cathodes at different potentials will need a separate heater battery or transformer winding each, and (2) there is a potential gradient along the cathode by virtue of the voltage dropped across it which means that it is a moot point what the 'cathode potential' is.

INDIRECTLY HEATED CATHODE

Here the cathode is a nickel tube coated with highly emissive material and filled with a refractory insulating material such as alumina; threaded through it is the heating element, a hairpin-shaped piece of tungsten wire (*Figure 8.2*).

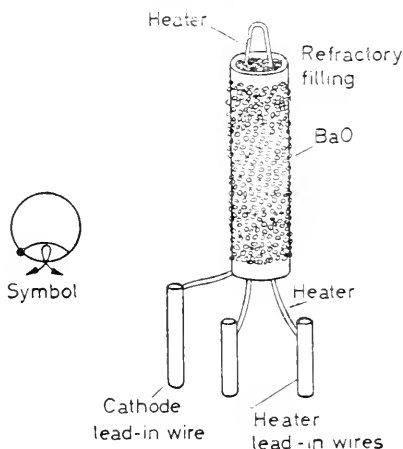


Figure 8.2

Such a cathode has a large thermal capacity and may be heated by a.c. or d.c. The cathode and the heater are insulated from each other, so that a number of valves with different cathode potentials may be fed from a common heater supply; furthermore the cathode surface is equipotential. On the whole the indirectly heated cathode is the more versatile device, and valves possessing them are much more common in electrobiological work.

OTHER ELECTRODES

Around the cathode is the anode, usually approximately cylindrical and made of nickel. Between them, according to the type of valve, there may be up to three 'grids'; these are concentric helices of nickel wire. There is also another object in the valve called a 'getter', but this is concerned with the manufacturing process and not with the operation of the valve.

In biological amplifiers valves are often 'under-run'—that is, the cathode temperature is arranged to be lower than that intended by the manufacturers. As we shall see later this is dictated by necessity; as a general rule it should be avoided, and care should be taken that the heater voltage is correct and

HARD VALVES

that some 'high tension' is always applied to the valve. If the HT is absent or if the heater voltage is too low then the cathode is liable to be 'poisoned' and its emission fail. If the HT is too high, i.e. exceeds the upper limit set by the manufacturers, the power dissipated in the valve will be excessive and the anode will overheat; when this happens there is a risk of the liberation of absorbed gas and consequent cathode destruction by ionic bombardment. If the heater voltage is excessive the cathode will overheat and the valve will fail from premature evaporation of the barium and strontium oxides deposited on it.

DIODE

If a hard diode current be measured as a function of the potential applied to the anode (*Figure 8.3*), the resulting relationship is of the form shown in *Figure 8.4*. The portions of the curve *AB*, *BC*, and *CD* represent three distinct regimes:

AB. Here the anode is negative with respect to the cathode, and electrons

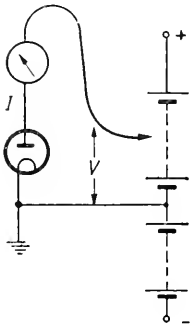


Figure 8.3

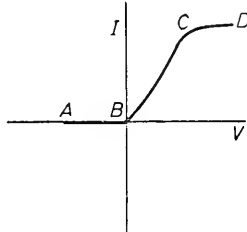


Figure 8.4

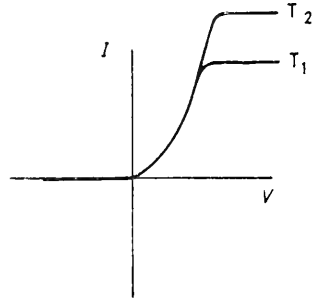


Figure 8.5

emitted by the cathode are turned back by the repulsive field due to the anode: no current flows.

BC. Anode moderately positive with respect to the cathode. Here the average rate at which electrons break clear of the cathode depends upon the resultant of the repulsive effect due to the negative space charge of electrons already emitted, and the attractive effect of the anode. The current is said to be 'space charge limited' and is proportional to $V^{3/2}$. This is called Child's law, or merely the 'three halves power law'. At *C* electrons are being removed from the cathode region to the anode as fast as they can be emitted, and the space charge disappears. If V be further increased there can be no corresponding increase in current because it cannot exceed the electron emission rate given by the Richardson equation.

CD. Between *C* and *D* the current is said to be 'temperature limited' and the diode is 'saturated'. If we are working near the point *C* between *C* and *D*, reversion to space charge limited working can be had by raising the cathode temperature, for the curve then flattens out at a higher value of I (*Figure 8.5*). A saturated diode is important because it is an approximation to a constant-current generator, provided the anode voltage is kept high enough. Special valves are made for this purpose, since ordinary valves are used under space

TRIODE

charge limited conditions and cannot be made to saturate unless: (1) the cathode temperature is reduced below that intended, which we have seen to be bad; or (2) the anode voltage is excessive, which is also bad.

TRIODE

When Lee de Forest put a grid between the anode and cathode of a hard diode he produced the key device from which the whole technology of electronics has grown up, the triode valve (*Figure 8.6*). Not only does this



Figure 8.6

device amplify, which facilitates the detection of weak effects, but it does so whilst imposing a load on the effect which is for many applications quite negligible—that is, its input resistance can be made very high. This distinguishes it from the other amplifiers used in electronics, the transducer and the transistor. At present its field of application is much wider than that of the transducer and transistor, though it is neither as robust as the former nor as small and economical in power consumption as the latter.

If a triode be set up as shown in *Figure 8.7*, and the anode current plotted as a function of anode voltage for various values of negative grid bias, then a series of diode-like characteristic curves are produced, each one displaced further to the right as the grid bias is increased (*Figure 8.8*). The mechanism

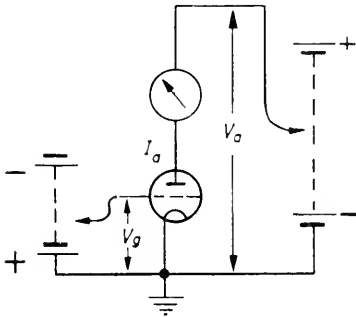


Figure 8.7

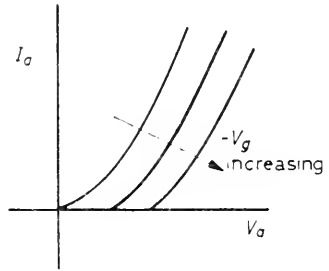


Figure 8.8

of this is that the negativity of the grid reinforces the effect of the space charge in reducing the anode current, and a higher anode voltage is necessary to achieve the same anode current. *Figure 8.8* is called the ‘anode characteristic’ of the valve.

There are three fundamental ‘valve parameters’ which describe the performance of a triode. They are: (1) μ , the amplification factor; (2) g_m , the mutual conductance; and (3) r_a , the anode incremental resistance.

μ is defined as $(\delta V_a / \delta V_g)_{I_a}$ and is the most nearly constant of the valve parameters, both with regard to the age of the valve and to the particular voltages and currents at which it is worked. μ is a pure negative number.

HARD VALVES

r_a is defined as $(\delta V_a / \delta I_a)_{V_g}$ and is given by the reciprocal of the slopes of the curves in *Figure 8.8*. Clearly it is moderately constant if I_a is not too small. It rises as the valve wears out. r_a is measured in ohms.

The third parameter g_m can be found by re-plotting the information of *Figure 8.8* in terms of the anode current against grid voltage for various anode voltages. It then appears as in *Figure 8.9*, which is called the grid

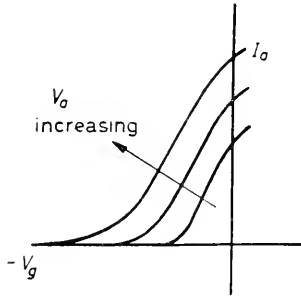


Figure 8.9

characteristic of the valve. The mutual conductance is $(\delta I_a / \delta V_g)_{V_a}$ and is given by the slopes of the curves in the grid characteristic. It is moderately constant if I_a is not too large or too small; it falls as the valve wears out. Its units are amperes per volt, but it is usually measured in milliamperes per volt.

It is not necessary to plot out the grid characteristic to find g_m . It can all be done from the anode characteristic—usually supplied by the valve manufacturers—for if

$$\mu = \left[\frac{\delta V_a}{\delta V_g} \right] \quad \text{and} \quad r_a = \left[\frac{\delta V_a}{\delta I_a} \right]$$

then

$$\frac{\mu}{r_a} = \frac{\left[\frac{\delta V_a}{\delta V_g} \right]}{\left[\frac{\delta V_a}{\delta I_a} \right]} = \left[\frac{\delta I_a}{V \delta V_g} \right] = g_m$$

As we have seen, μ and r_a can be read off the anode characteristic; g_m is given by their ratio.

In general it is sufficiently true to say that, because the triode valve is operated with the grid biased negatively with respect to the cathode the electrons will be obliged to pass between the interstices of the grid on their way to the anode—being repelled away from the grid wires—and that therefore no current can flow into or out of the grid. This is the same as saying that the resistance ‘looking in’ at the grid of the valve is infinite. We say that the valve is ‘voltage operated’ rather than ‘power operated’.

SIMPLE AMPLIFICATION WITH THE TRIODE

To get amplification or ‘gain’ with a triode valve we cause the anode current to flow through a load resistance. This resistance may be an actual load

SIMPLE AMPLIFICATION WITH THE TRIODE

such as a penwriter or a loudspeaker, or it may be included so that fluctuations in the anode current caused by a signal at the grid are converted to fluctuations of potential difference across the load so that these may be passed on to operate another valve. The former case is called power amplification and the latter voltage amplification. We consider voltage amplification first.

Simple voltage amplifier

In *Figure 8.10* we see a triode valve whose anode potential is derived from a HT supply—in this case a battery—via a load resistance R . The grid is given a mean negative bias by another battery as shown, and a signal V_{sig} is superposed on this mean bias by placing the generator, which we regard

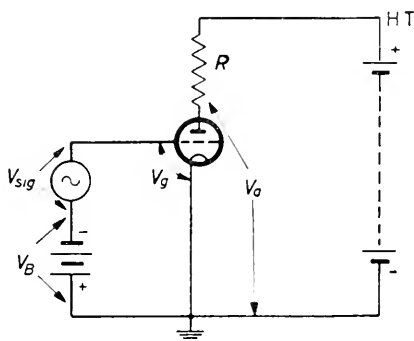


Figure 8.10

as the source of the signal, in series with it. V_g is thus $-V_B + V_{\text{sig}}$. A mean current flows down the load resistance and through the valve, producing a voltage drop across the load so that the anode potential is less than the HT voltage. The bias and anode voltages are chosen so that the valve operates on a linear part of the anode and grid characteristics, so that the parameters g_m , r_a and μ are as nearly as possible constant.

At first sight it might appear that as $\delta V_g = V_{\text{sig}}$, $\delta I_a = g_m V_{\text{sig}}$ and therefore $\delta V_a = R g_m V_{\text{sig}}$. Then the amplification would be $\delta V_a / V_{\text{sig}} = g_m R$. In fact this is not so, for if V_a changes, $\delta I_a / \delta V_g \neq g_m$. In words, suppose the grid be made a certain amount less negative. The anode current increases but there is an increased voltage drop across the load and the anode becomes less positive. As a result the anode current does not increase as much as it would have done had the anode potential been fixed. The actual amplification can be found either analytically or graphically. The answer can be found in a few seconds from the analytic solution for an existing design. In creating a new design the graphical approach has to be adopted anyway because it is the quickest way of finding the required load resistance and bias.

Analytic expression of amplification

$$\text{If} \quad g_m = \left[\frac{\delta I_a}{\delta V_g} \right]_{V_a} \quad \text{and} \quad r_a = \left[\frac{\delta V_a}{\delta I_a} \right]_{V_g}$$

$$\text{then} \quad \delta I_a = g_m \delta V_g + \frac{1}{r_a} \delta V_a$$

HARD VALVES

but

$$\begin{aligned}\delta V_a &= -R\delta I_a \\ \therefore \delta I_a \left(1 + \frac{R}{r_a}\right) &= g_m \delta V_g \\ \delta I_a &= \frac{g_m \delta V_g}{1 + \frac{R}{r_a}}\end{aligned}$$

The amplification is

$$\frac{\delta V_a}{\delta V_g} = \frac{\delta I_a R}{\delta V_g} = \frac{g_m R}{1 + \frac{R}{r_a}}$$

The expression may be left in this form, or, multiplying through by r_a

$$\text{amplification} = \frac{g_m r_a R}{r_a + R} = \frac{\mu R}{R + r_a}$$

A typical value for a triode valve would be 40.

The expression tends to an upper limit equal to μ when $R \gg r_a$, hence the term amplification factor. It should, however, be borne in mind that if, in an effort to achieve high amplification, R is made very large, then the anode voltage and current will fall and so will μ , and the expected increase will not be secured. μ can only then be restored by increasing the HT voltage. In point of fact there are better methods of getting high gain than squeezing the last drop of amplification out of a triode by the employment of very large high tension. It is best to be content with the gain available from a high tension supply which is convenient.

Graphical design procedure

Choose a valve of high μ and from the makers' catalogue make a tracing of the anode characteristic. Our problem now is to find the 'working region'. This is bounded by:

- (1) A rectangular hyperbola which represents the maximum power which may be dissipated at the anode of the valve (lest it get too hot).
- (2) A vertical line through the voltage which is the maximum which may be applied to the anode (lest it liberate gas).
- (3) A horizontal line which represents the smallest current below which curvature of the characteristic curves sets in (lest the output cease to be a linear function of the input).
- (4) The characteristic curve—guess it if it is not there—corresponding to a grid voltage of about $-\frac{1}{4}$ V (lest there flow grid current).
- (5) A horizontal line which corresponds to the maximum emission current at which the cathode is rated.

Then the area enclosed by these five lines is the working region. Draw a dot in what you guess to be about the middle. Then this point is a first shot at the 'working point' for the valve, and the mean anode current, anode voltage and grid bias may be read off.

On the voltage axis mark off the proposed value of high tension and draw a straight line through the high tension point and the working point. Then

SIMPLE AMPLIFICATION WITH THE TRIODE

the reciprocal of the slope of this line gives the value of anode load required in order that the valve anode voltage and anode current shall be correct when the bias has the proposed value. This line is called the load line (*Figure 8.11*). A signal applied to the grid adds to and subtracts from the standing bias, altering the grid voltage, and sliding the working point along the load line. The amplification of the proposed circuit may be found by seeing how

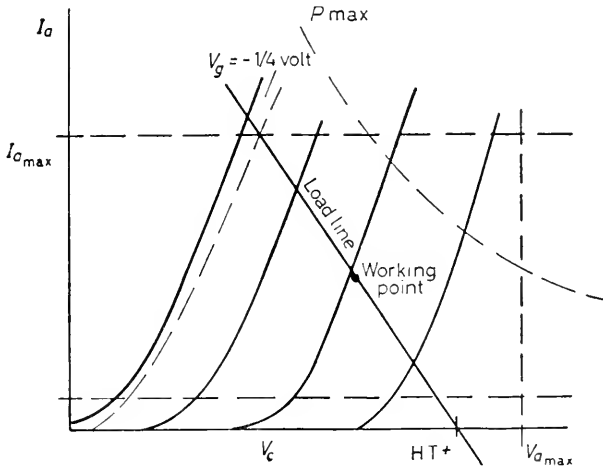


Figure 8.11

much the change of anode voltage is for a unit change of grid voltage, and the largest signal the valve can accept is read off as the largest change of grid voltage in either direction which just does not carry the working point outside the working region.

We now modify the design in the light of particular requirements. If the valve is a 'low level' amplifier—that is, the input signals are small (of the order of microvolts)—it will be shown later that it is helpful to use a low anode potential and convenient to have a low high tension voltage because it is best derived from a battery. Because the working point is not moved very far by the signal it is quite safe to arrange it near the edge of the working region, so choose one at the bottom left-hand corner. Mark the new HT point—HT batteries are commonly 60, 90 or 120 V—and read off the load and bias required, and the gain. Do not worry if the latter is not very large.

In a 'medium level' amplifier—input signal of the order of millivolts—concentrate on securing the maximum possible gain. Choose a high tension as large as convenient and the flattest load line which will allow a working point somewhere near the middle of the bottom of the working region. Read off the load, and the gain which would be obtained and the bias required.

With a 'high level' amplifier—input signal of the order of volts—the prime requirement is for the valve to be able to accept the large input signal without the working point passing outside the working region; a low gain is to be expected. The provisional design is probably about right for a high level amplifier, but see whether by juggling with the load line, the HT and the mean working point the valve can be made to accept a larger input. If when

HARD VALVES

conditions seem to be optimum the signal handling capacity is still not good enough, begin the design again with a different valve of lower μ .

Simple direct-coupled power amplifier

To achieve power amplification the circuit diagram is similar to the voltage amplifier (*Figure 8.12*) but the component values are rather different. A typical value of load resistance for a voltage amplifier might be 50,000 ohms,

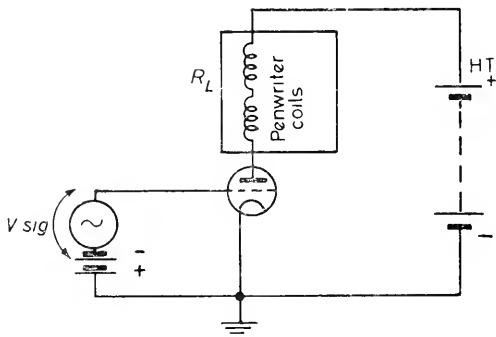


Figure 8.12

while the r_a of the valve could be 10,000 ohms. For a power amplifier we want the maximum power transferred to the load and this happens when $R_L = r_a$, for the valve behaves like a constant voltage generator of e.m.f. $\mu\delta V_g$ in series with a resistance r_a (*Figure 8.13*). $R_L = r_a$ may not be very easy to arrange. If the frequency response of the system has to extend down to zero frequency the load must be placed directly in series with the valve as in *Figure 8.12*. The problem is usually to get the load resistance high enough and the r_a low enough. The load might typically be the coils of a

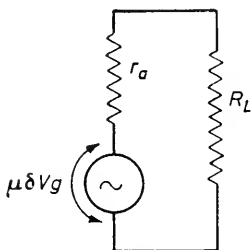


Figure 8.13

penwriter or the magnetic deflection system for a cathode ray tube, and high resistance is only usefully obtained by winding the coils with enormous numbers of turns of very fine wire, a procedure which leads to expense and fragility. The necessary low r_a can only be obtained by using a valve which tends to be large, since its electrode structure may be regarded as a number of valves of more ordinary r_a connected in parallel.

In practice we may have to accept a considerable mismatch— R_L several times less than r_a —and this is unfortunate on two counts; (1) because of the mismatch the power transfer conditions are less than optimum; (2) it is not possible to apply as large an input to the valve as would otherwise be

SIMPLE AMPLIFICATION WITH THE TRIODE

the case. That is, the valve is overloaded more easily*. This may easily be seen from the working region of the power amplifier valve anode characteristic (*Figure 8.14*). Clearly, when the load is small (load line steep) the maximum

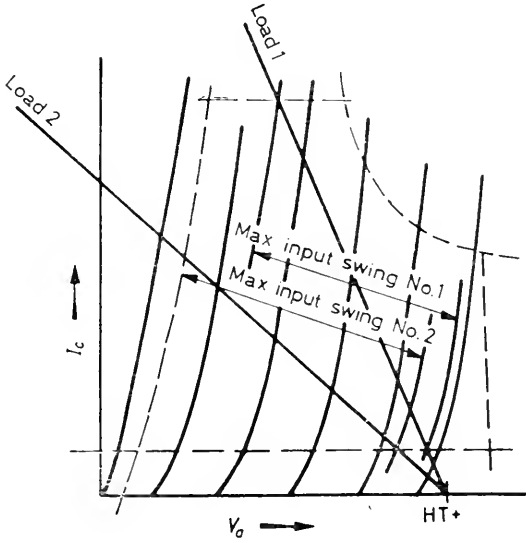


Figure 8.14

possible input 'swing' δV_a is much less than for the flatter load line representing the higher load.

Simple transformer-coupled power amplifier

If the response of the system does not have to extend down to zero frequency, as for example in sound reproduction, design is much easier, for a transformer may be used to match up any convenient valve and load. The

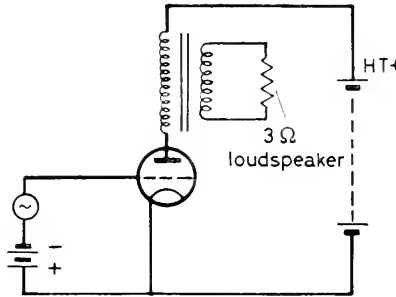


Figure 8.15

resistance of a loudspeaker is commonly as low as 3Ω . An appropriate valve might have $r_a = 5,000 \Omega$. The two can be matched for approximate maximum power transfer by using a transformer of turns ratio $(5,000/3)^{1/2}$ (*Figure 8.15*).

* An unfortunate but accepted piece of jargon. Is nothing to do with excessive values of load. 'Overloading' simply means an excessive input that takes the working point outside the working region.

HARD VALVES

The presence of the transformer modifies slightly the procedure for graphical design, for we have now two kinds of load. (1) The 'static load' or 'd.c. load', which is the load relevant to the determination of the mean working point as defined by the steady—or zero frequency—bias, anode voltage and current. The static load is the resistance of the transformer primary winding, which is usually negligible. (2) The 'dynamic load' or 'a.c. load' which is the load as seen by the load when handling signals, and which is given by R_L . The diagram for graphical design is drawn in the following manner.

Choose a mean working point. Drop a vertical straight line through it to the voltage axis. This is the d.c. load line and the HT required may be

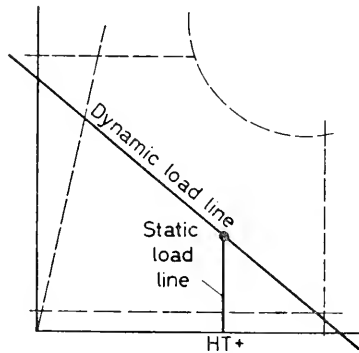


Figure 8.16

read off. Draw the dynamic load through the working point, having the appropriate slope (Figure 8.16).

Choice of transformer ratio—For a given input signal the maximum power is transferred to the load when the optimum transfer conditions are met. In practice these conditions are deliberately not met; let us see why this is.

If the load resistance reflected into the valve circuit is other than r_a the system is mismatched, and more voltage swing will be necessary at the valve grid to develop the same power in the load. This is usually a minor problem and can be dealt with by a bit more amplification at an earlier stage. A much more serious difficulty is to get a load condition such that the greatest voltage swing may be applied to the valve grid without taking the working point outside the working region. The 'optimum load' is the load into which the valve will deliver the maximum signal power without distortion. We saw that with the direct coupled power amplifier low values of load were unfavourable from the point of view of possible input swing. For a triode it can be shown that the optimum load is *twice* r_a , for the loss of output power due to the mismatch is overshadowed by the possibility of increasing the input. We shall not prove it rigorously, but the reader may care to experiment with some load lines on an actual published anode characteristic. The output power and efficiency of a power amplifier may be read off as follows. In Figure 8.17, if the grid swings between V_{g1} and V_{g2} , the anode current swings between E and F and the anode voltage between G and H . Now these are peak-to-peak values. For a sine wave input the R.M.S. voltage

TETRODE

swing = $(GH)/(2\sqrt{2})$ and the R.M.S. current swing = $(EF)/(2\sqrt{2})$. The R.M.S. output power is therefore

$$\frac{\text{Area of rectangle } ABCD}{8}$$

Similarly for a square wave input

$$\text{R.M.S. output} = \frac{\text{Area of } ABCD}{4}$$

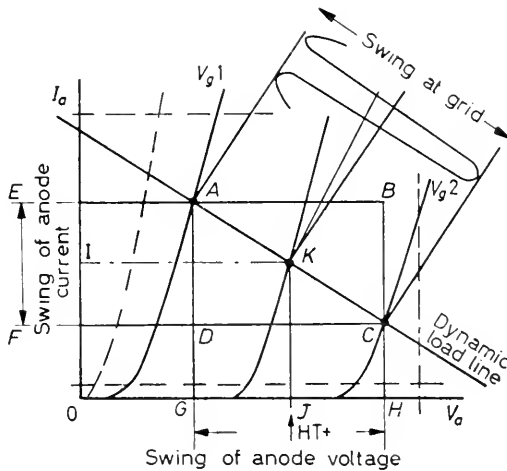


Figure 8.17

The mean input power to the stage = $HT \times \text{Mean anode current}$
 = Area of rectangle $OIKJ$

Efficiency of power amplifier =

$$\frac{\text{Area } ABCD}{8 \times \text{area } OIKJ} \times 100 \text{ per cent for sine waves}$$

or
$$\frac{\text{Area } ABCD}{4 \times \text{area } OIKJ} \times 100 \text{ per cent for square waves}$$

Notice that with transformer coupling the anode voltage swings above the HT. The same construction may be used to find the power output and efficiency of direct coupled power amplifiers.

TETRODE

The simple triode valve amplifier has two important disadvantages; one we have already mentioned—its voltage amplification is not very great, being $(g_m R)/(1 + R/r_a)$ and not $g_m R$ as we might hope. Matters will be improved if we can find some way of increasing r_a ; looked at another way, it will help if we can reduce the dependence of the anode current on the anode voltage.

Miller effect

The other disadvantage is perhaps rather unexpected—it is that the input capacitance is high; the valve seems, looking in at the grid, to have an appreciable capacitance between grid and cathode. One of the most important reasons for using valves is that the input impedance is high, which means that not only the input resistance but also the input reactance is large. This appreciable virtual input capacitance of triodes is called ‘Miller effect’ and arises in the following way.

Suppose we have a generator, a hypothetical ‘charge-meter’, perhaps a ballistic galvanometer, a capacitance and a ‘black box’, all in series as in *Figure 8.18*, and suppose initially that the black box contains nothing but a

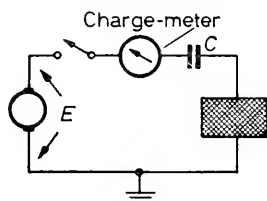


Figure 8.18

piece of wire joining the two terminals. Then on switching on the generator of e.m.f. E a charge Q_1 is displaced round the circuit which we measure on our charge-meter, and we construe that the capacitance has a value $C = Q/E$. Now let the black box contain a device which, when the generator is switched on, immediately makes the right-hand plate of the capacitance AE volts negative to earth. Then the actual voltage across the capacitance $= E(1 + A)$ V and the charge displaced will be $Q_2 = CE(1 + A)$ instead of $Q_1 = CE$. The circuit therefore behaves as if the black box contained once more just the piece of wire, but as if the capacitance were not C but $C(1 + A)$.

This is precisely the state of affairs with the triode amplifier. The grid has a capacitance C_{gk} to the cathode and C_{ga} to the anode. A unit change of grid potential produces a potential change at the anode of $-A$, where A is the gain. The apparent input capacitance is therefore not $(C_{gk} + C_{ga})$ but $C_{gk} + (1 + A)C_{ga}$.

A triode might have $C_{gk} = 5 \mu\mu\text{F}$ and $C_{ga} = 3 \mu\mu\text{F}$. If the input capacitance were only $8 \mu\mu\text{F}$ this would not be very serious, as the reactance at 10,000 c/s would be

$$\frac{1}{2\pi \times 10,000 \times 8 \times 10^{-12}}$$

which is about 2 megohms. However, if $A = 50$, the input capacitance is $5 \mu\mu\text{F} + 3 \times 51 \mu\mu\text{F} \approx 160 \mu\mu\text{F}$. At 10,000 cycles the reactance of this is

$$\frac{1}{2 \times 10,000 \times 160 \times 10^{-12}}$$

which is only about 100,000 Ω , and is for many purposes intolerably low.

In the tetrode a second grid is placed between the ‘control grid’ and the anode, called the ‘screen’ grid, and in use this is maintained at a fixed positive

TETRODE

potential (Figure 8.19). This grid forms an electrostatic screen between the anode and the control grid (which enormously reduces C_{ga}) and also screens the cathode region from the potential changes at the anode. The rate at which electrons leave the cathode region, or the 'cathode current', is determined almost entirely by the screen potential, but due to the open structure of the screen grid most of the electrons pass through it and strike the anode;

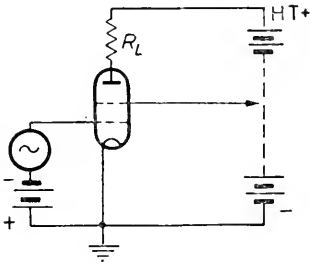


Figure 8.19

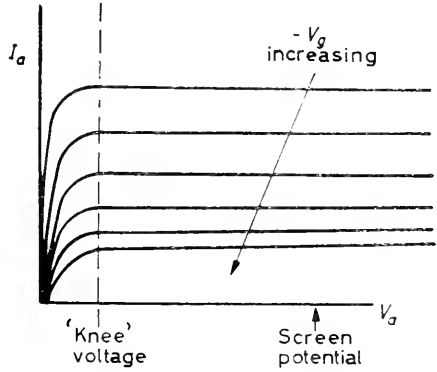


Figure 8.20

a small proportion hit the screen and constitute 'screen current'. This is the rationale of the tetrode and the anode characteristic for which one might hope has the form of Figure 8.20. Some anode potential is necessary to get anode current, for clearly if there were none at all, all the cathode current would pass to the screen. However, once the anode voltage exceeds a critical 'knee voltage', substantially all the cathode current passes through the screen grid to the anode, giving the desirable characteristics already enumerated.

In point of fact the anode characteristic has the curious form of Figure 8.21.

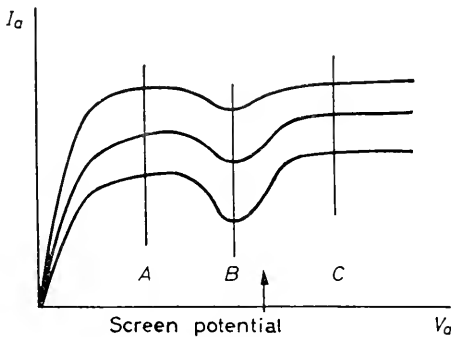


Figure 8.21

The curves have 'tetraode kink'; this may be explained as follows: When the anode potential is at A the electrons pass through the screen and encounter a retarding field because the anode is negative with respect to the screen. Under the influence of this field they are slowed up and finally strike the anode

HARD VALVES

rather gently. When the anode potential is at *B* the retarding field is much less and the electrons hit the anode hard, knocking out 'secondary electrons'. These travel back to the screen because it is more positive than the anode, so that the net anode current is less than expected—hence the kink. At *C* the secondary emission is even stronger, but because the anode is now more positive than the screen the secondary electrons fall back on the anode. There is a device called the dynatron which makes use of tetrode kink: the author has never found it being used in biological work so it will not be described. On the whole tetrode kink is a bad thing, and to get rid of it further elaboration is needed. The valves which overcome it are the pentode and the beam tetrode.

PENTODE

In the pentode yet another grid is included, between screen grid and anode. This is maintained at a fixed negative potential with respect to the screen,

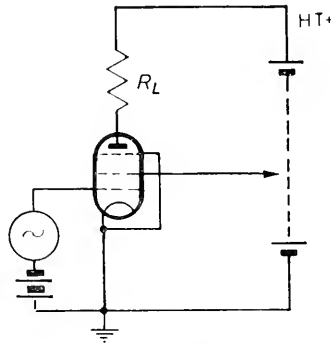


Figure 8.22

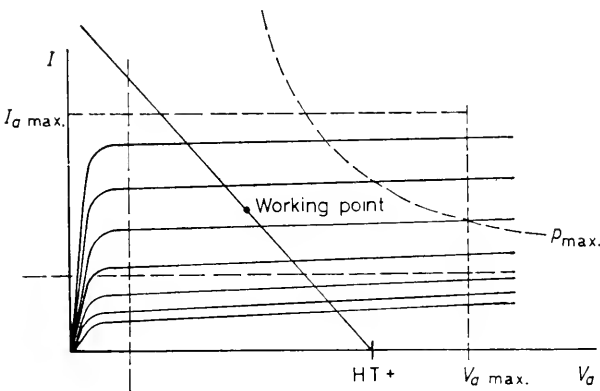


Figure 8.23

usually by connecting it to the cathode (Figure 8.22). It is called the suppressor because its function is to suppress the flow of secondary electrons from anode to screen by screening the anode from the screen. There is then no tendency

BEAM TETRODE

for electrons knocked out of the anode to do anything but fall back on to it again.

With the pentode the anode characteristic is as in *Figure 8.20* and voltage amplifiers may be designed graphically in the same manner as for triodes (*Figure 8.23*). Since at the bottom of the diagram the curves tend to close up, the bottom edge of the working region is bounded by an arbitrary level at which the cramping becomes intolerable. The left-hand edge is bounded by a voltage just above the knee voltage. The remainder of the working region is bounded by the hyperbola of maximum anode dissipation, and by the maximum permissible anode voltage and current, as in the triode. The amplification may be taken as $g_m R$ and is typically 150 times. Pentodes may also be used for power amplification, and will deliver more power for a given mean anode voltage and current than will a triode, as explained in *Figure 8.24*. Due to the cramping of the curves at the bottom, pentodes,

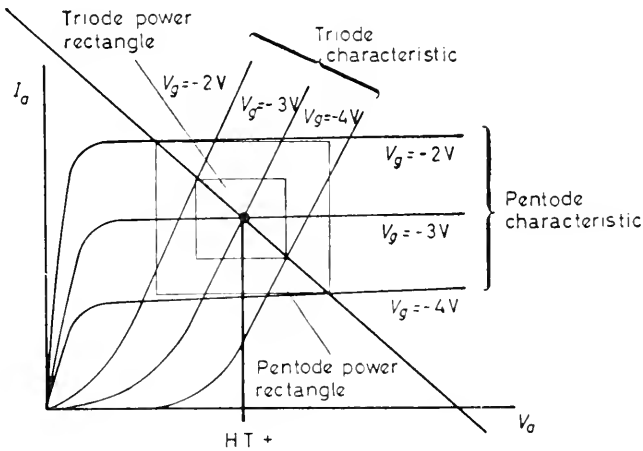


Figure 8.24

whether high-level voltage or power-amplifying, tend to cause more distortion than triodes. The optimum load for power pentodes is only about $1/5$ to $1/10$ of r_a ; but as r_a is so high, of the order of hundreds of thousands of ohms, the optimum load usually works out higher than for a triode.

BEAM TETRODE

A beam tetrode is a carefully made—and therefore expensive—valve suitable only for power amplification (*Figure 8.25*). The screen grid is wound in the shadow of the control grid (as illuminated by the cathode) as a result of which there is a clear run for the ‘beaming’ of electrons on to the anode, and incidentally the screen current is extremely small. ‘Beam forming plates’ at cathode potential are included which constrain the beam to strike the anode at right angles. Tetrode kink is prevented because: (1) secondary electrons released by primary incident electrons which strike normally are of minimal energy; and (2) *because* the valve is a power valve and *because* the primary electrons are beamed the electron density is so great that the

HARD VALVES

concomitant space charge is sufficient to turn back secondarily emitted electrons. The anode characteristics of beam tetrodes and power pentodes are similar and design procedure is the same for each.

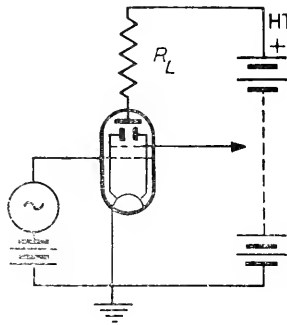


Figure 8.25

Warning—The anode circuit of power pentode or beam tetrode valves must never be broken whilst the screen potential is applied. If this is done all the cathode current will go to the screen, which is of insufficient thermal capacity to withstand the resulting temperature rise. It is extremely likely that the screen will melt and the valve be destroyed.

Equivalent circuits of valve amplifying stages

The voltage gain expression for a simple triode amplifier $\mu R / (R + r_a)$, suggests the equivalent circuit for the stage introduced in Figure 8.13, namely, a $\mu \delta V_g$ volt generator in series with a resistance r_a , feeding the load R_L . Similarly, Figure 8.26a shows an appropriate equivalent circuit for a pentode

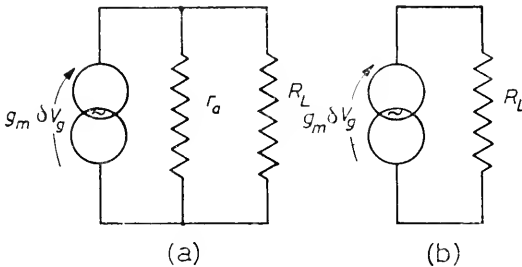


Figure 8.26

or beam tetrode stage. Since r_a is usually so much greater than R_L , its contribution to the two resistances in parallel is generally negligible and the valve may be regarded practically as a generator of constant current $g_m \delta V_g$ feeding the load R_L (Figure 8.26b).

INTERSTAGE COUPLING

Where the gain of the simple voltage amplifier is insufficient we achieve the requisite amplification in several 'stages' by coupling the output of one valve to the input of the next. Again, when the input swing required by a power amplifying valve in order that the load be fully operated is greater than the available source of signal can provide we interpose a voltage amplifying stage. In both cases the problem arises as to how to couple adjacent valves.

There are two broad divisions in the ways in which this can be carried out, depending on the frequency response required. If the response has to extend down to zero frequency the stages have to be 'direct coupled', but if it is possible to allot a lower frequency limit to the desired response the stages may be 'a.c. coupled'. As valves are voltage-operated devices it would be more logical if the term could be amended to 'a.v. coupled', but the notion of the a.c. coupled amplifier is well established and is likely to remain. As a.c. coupled amplifiers are much easier to design and use it would be foolish to use a direct coupled apparatus for, say, action potential recording from a nerve, where a frequency response extending down to 100 c/s would be ample. On the other hand if one is interested in the absolute value of a membrane potential one has no alternative but to use direct coupling*.

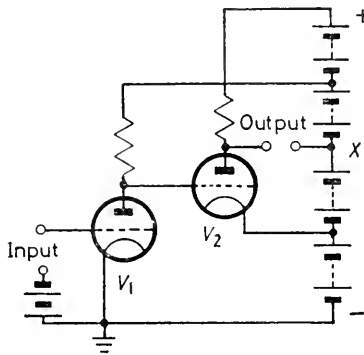


Figure 9.1

DIRECT COUPLING METHODS

Climbing amplifier

By having an HT battery provided with a very large number of taps the grid of the second valve may be connected directly to the anode of the first, in the manner shown in *Figure 9.1*, which shows two voltage amplifying stages.

* Except one, a special kind of amplifier called a carrier amplifier, with which we deal in Part IV.

INTERSTAGE COUPLING

The tap at X is chosen to be equal to the mean anode potential of V_2 so that there is no difference of potential between the output terminals when there is no difference of potential at the input. Climbing amplifiers work well, but have at least three disadvantages when more than about two stages are used:

(1) An inconveniently large total HT voltage is required.

(2) If earth is connected to the amplifier as shown in *Figure 9.1*, then the output is at a high positive potential with respect to earth. If the position of the earth connection is moved to, say, the point X , then the input terminals are inconveniently negative with respect to earth.

(3) If a common heater supply is used for all valves it may be difficult to arrange that the maximum allowable heater-cathode voltage for the valves is not exceeded, leading to breakdown of the alumina insulation.

All three difficulties are overcome if the amplifier be prevented from climbing, that is, if all the valves are operated with their cathodes at approximately earth potential. To do this we have to transfer the signal from valve to valve with a fall in voltage from the considerably positive anode of valve 1 to the slightly negative grid of valve 2.

Coupling battery

One way of doing this is with a battery. If the anode potential of V_1 in *Figure 9.2* is $+75$ V, and the bias required by V_2 is -3 V, then the correct

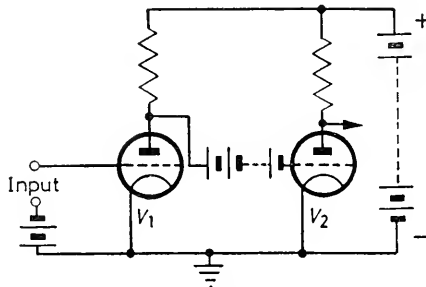


Figure 9.2

conditions are obtained when the e.m.f. of the coupling battery is 78 V.

The coupling battery does not have to supply any current and so can comprise small cells of the deaf-aid type. It is a bad thing to use a physically large coupling battery as the high frequency performance of the amplifier will be spoiled by the capacitance of the battery to earth. Deaf-aid batteries do not possess taps, so the design of battery-coupled stages is restricted by the available battery voltages. As the smallest and most convenient batteries can be made up from units giving multiples of 15 V, the mean anode voltage finally chosen in a design may have to be altered $\pm 7\frac{1}{2}$ V from the provisional value. This is not usually a very serious matter. A disadvantage with coupling batteries is that before they are many months old their voltage becomes rather erratic, so that spurious signals are generated in the amplifier and the batteries have to be replaced. With a large number of stages this is a tedious and expensive business.

Potential divider

For this method (Figure 9.3) it is necessary to have a negative supply voltage, but as one source can serve a whole amplifier it is much more economical in batteries than the coupling-battery method. It is clear that if the ratio of R_1 to R_2 is correctly chosen the grid potential of V_2 may be arranged to

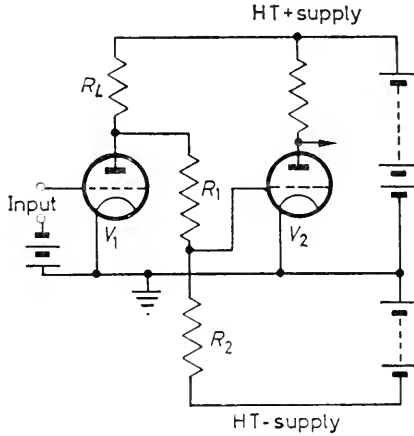


Figure 9.3

come out at the right bias. The fraction of the signal appearing at V_1 anode which is transferred to V_2 grid is $R_2/(R_1 + R_2)$, so that R_2 should be large; if R_2 is large, the negative supply must be large in order that V_2 grid potential be correct. R_1 and R_2 alter the working conditions of V_1 , for V_1 now has an effective load of $R_1 + R_2$ in parallel with R_L —thus reducing the gain—and effectively a reduced HT. Exact design becomes complicated but fortunately is not necessary, for it is clear that the reduction of V_1 gain and HT is minimal when $R_1 + R_2$ is large compared with R_L . We therefore arrange that this is so, then V_1 working conditions may be regarded as substantially unaffected.

Suppose the anode potential of V_1 is provisionally $+60$ V, the negative supply is -100 V, $R_L = 50,000 \Omega$, and the bias required by $V_2 = -2$ V.

Then
$$\frac{R_1}{R_2} = \frac{V_{R_1}}{V_{R_2}} = \frac{62}{98}$$

If we make $R_1 = 620,000 \Omega$ and $R_2 = 980,000 \Omega$, then $R_1 + R_2 = 1.6 \text{ M}\Omega$ which is certainly $\gg R_L$. The fraction of the signal transferred by the coupling will be $98/(62 + 98) = 0.6$. This is rather a serious loss. Suppose it is found that V_1 will give almost as much gain if the mean anode potential is dropped to $+30$ V.

Then $R_1/R_2 = 32/98$. Suitable values would be $320,000 \Omega$ and $980,000 \Omega$, for $R_1 + R_2 = 1.3 \text{ M}\Omega$ is still $\gg R_L$. The fraction of signal transferred will be $98/(98 + 32) = 0.755$, a considerable improvement. A still better performance will result if the circuit be re-designed for a negative supply of, say, -150 V.

The potential divider coupling method is cheap, simple and reliable. It

INTERSTAGE COUPLING

is also effective within the limitations already outlined. As it works best when V_1 anode potential is low it is particularly suited to the coupling of low-level amplifier stages.

'Neon lamp' coupling

This makes use of the ability of a cold cathode diode to maintain across itself a rather constant potential difference; clearly as the potential at A in *Figure 9.4* fluctuates the current in the ballast resistor R_b fluctuates also and

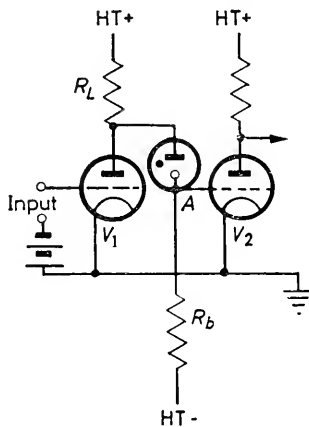


Figure 9.4

this current flows through the lamp, so that the constant voltage properties are essential. The advantage of the method is that it is almost as cheap and reliable as the potential divider method, but, like the coupling battery technique, provides almost no reduction in signal transferred.

In order that the performance of V_1 be not degraded by the coupling, R_b must be $\gg R_L$, which implies that the lamp current must be much smaller than V_1 anode current. About $100 \mu\text{A}$ is typical. Voltage reference tubes and small neon indicator bulbs will work satisfactorily at this current, but stabilizer tubes on the whole will not. Design procedure is therefore: having fixed a provisional mean anode potential V_a for V_1 and grid bias $-V_g$ for V_2 , find the reference tube or indicator neon whose running voltage is the nearest to $(V_a + V_g)$. Take the makers' minimum current and divide it into the negative supply voltage. This gives with sufficient accuracy the value of R_b . Now re-design the first stage for an anode current of at least five times the neon current and for the corrected anode potential as dictated by the particular neon lamp chosen.

The negative supply must be at least more negative than the voltage differential of the reference tube or lamp, otherwise the latter will not strike. It ought to be considerably more negative than this in order that the glow shall not extinguish during large negative excursions of V_1 anode.

The field of application for neon coupling is in the high level stages of amplifiers for two reasons:

(1) the voltage across a soft diode is subject to small fluctuations and these must be swamped by typical signals;

A.C. COUPLING METHODS

(2) the minimum current of many voltage reference tubes is as high as 2 mA. A V_1 anode current of 10 mA is therefore required, an amount much too great for low-level stages.

A.C. COUPLING METHODS

There are two a.c. coupling methods: transformer coupling and resistance-capacitance coupling. Transformer coupling was *de rigueur* in the days when the amplification obtainable for valves was not very great and the voltage step-up obtainable with a transformer was welcome. Modern valves give very high gains, and transformers are bulky, expensive and complicate design; in consequence the intervalve coupling transformer is practically extinct.

Resistance capacitance coupling

The resistance and capacitance referred to are R_g and C in *Figure 9.5*. C is the coupling capacitance or 'blocking' capacitance, because it prevents the positive anode potential of V_1 from affecting the negative bias potential of

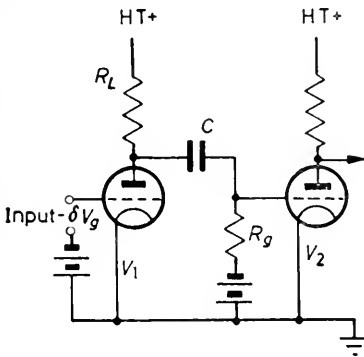


Figure 9.5

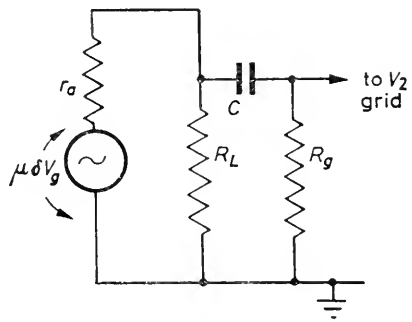


Figure 9.6

V_2 . The exact equivalent circuit of the device is shown in *Figure 9.6*. Since r_a and R_L are usually much smaller than R_g this is with sufficient accuracy reducible to *Figure 9.7*. If V_1 is a pentode the equivalent circuit is *Figure 9.8*,

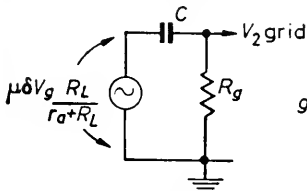


Figure 9.7

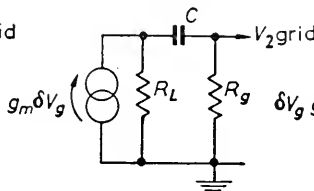


Figure 9.8

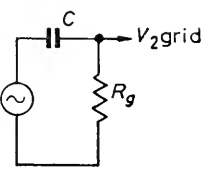


Figure 9.9

which, if $R_L \ll R_g$, reduces to *Figure 9.9*.

In both cases the coupling is seen to form a high-pass filter of time constant CR_g or turn-over frequency $\omega_c = 1/(CR_g)$. ω_c must therefore be below the lowest frequency the amplifier is required to transmit without significant loss

INTERSTAGE COUPLING

of amplification. A reasonable design procedure is as follows: Take the lowest frequency the amplifier is required to handle and divide it by the number of RC couplings there are going to be. This gives ω_c . Take the reciprocal, getting the minimum product CR_g for each coupling. There is usually a maximum value of R_g which the makers of V_2 will allow. It is called the maximum grid resistance and is given in the manufacturers' literature. Having found this, the economic minimum value of C follows. There is no point in making C larger, as 'microfarads cost money'.

Automatic grid bias—With RC coupled amplifiers it is possible to dispense with the bias batteries by using 'automatic grid bias', which is illustrated in *Figure 9.10*. The bottom of the grid resistance is taken to earth and instead

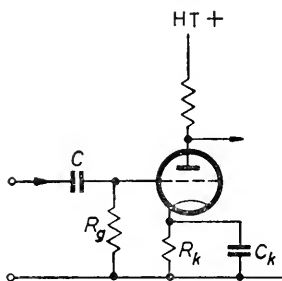


Figure 9.10

of making the grid slightly negative to earth we make the cathode slightly positive, by causing the cathode current to flow down a resistance. The value of resistance R_K required is simply

$$R_K = \frac{\text{Requisite bias}}{\text{Cathode current}}$$

Notice the use of the term cathode current rather than anode current. In *Figure 9.10* a triode is shown and the two are the same thing. If the valve had been a pentode or tetrode, the cathode current would have been the sum of the anode and screen currents. The screen current is usually small but it ought not to be forgotten. The cathode current fluctuates with the input signal and produces a fluctuating bias voltage across it. This is frequently undesirable; the fluctuation in bias may be 'ironed out' by connecting across the bias resistance a large capacitance C_K , the 'cathode bypass capacitance'. This must be chosen so that its reactance is always much lower than the resistance of R_K , even at the lowest frequency the amplifier is to transmit. Values of the order of 100 microfarads are often necessary. The biasing scheme is automatic in the sense that as the valve wears out and the anode current falls, the bias is automatically reduced and a measure of compensation obtained.

FREQUENCY RESPONSE OF INTERVALVE COUPLINGS

We have already seen that whilst the frequency response of a direct-coupled pair of valves extends down to zero, the response of an a.c. coupled pair is that of a simple high-pass filter, that is, the response curve is 3 dB's down at

FREQUENCY RESPONSE OF INTERVALVE COUPLINGS

$\omega = 1/(CR_g)$ and is asymptotic to a slope of 6 dB octave. At high frequencies the response of both types of coupling falls off, also at 6 dB octave, for a simple low-pass filter is made by V_2 grid- and stray-capacitance in conjunction with the effective internal resistance of the source driving it. Consider the a.c.

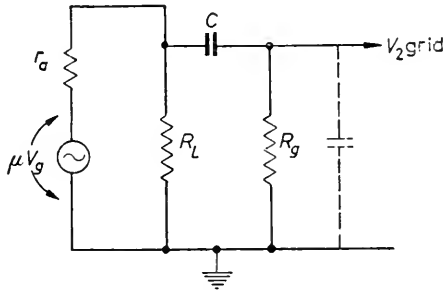


Figure 9.11

coupled case. The equivalent circuit for a triode V_1 is Figure 9.11, as we have seen. At high frequencies the reactance of C is negligible but the reactance of the shunt capacitance is not. This shunt capacitance is made up of V_2 input capacitance, $C_{gk} + (1 + A)C_{ga}$, plus C_S , the stray wiring capacitances to earth. R_g is large compared with R_L and may be neglected (since it is now virtually in parallel with it) so the equivalent circuit reduces to Figure 9.12.

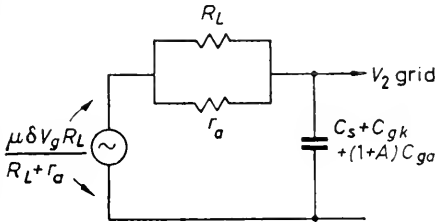


Figure 9.12

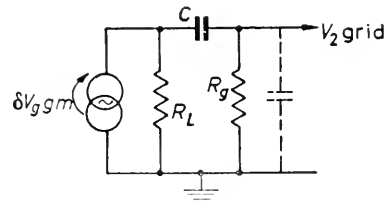


Figure 9.13

If V_1 is a pentode we have Figure 9.13 which reduces to Figure 9.14. It is clear that Figures 9.12 and 9.14 are simple low-pass filters whose performance may be calculated in the usual way.

Evidently the high frequency response is best when: (1) r_a and the load

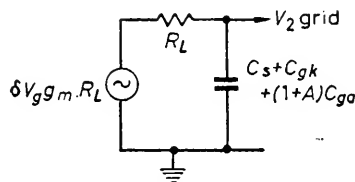


Figure 9.14

resistance of V_1 are low; (2) the input capacitance of V_2 and the stray capacitances are low. (1) is achieved by making V_1 a triode rather than a pentode. (2) is achieved by making V_2 a pentode, because pentode input

INTERSTAGE COUPLINGS

capacitance is low due to the substantial suppression of Miller effect. (2) is a much greater effect than (1), and as V_1 has itself to be coupled to a source possessing internal resistance amplifiers often employ pentodes throughout. With direct coupled amplifiers the equivalent circuits at high frequencies are similar, the climbing and neon-coupled types performing best. With coupling batteries the upper cut-off frequency is liable to be somewhat low because of the stray capacitance of the bulky battery to earth. The potential divider type requires rather more comment. Referring back to *Figure 9.3* but assuming V_1 and V_2 to be pentodes, the equivalent circuit at high frequencies is *Figure 9.15*. Sectioning the circuit along the dashed line and applying

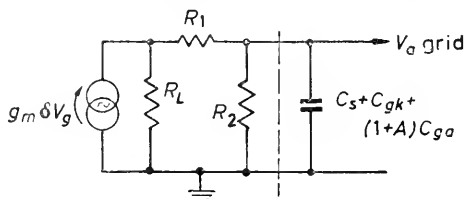


Figure 9.15

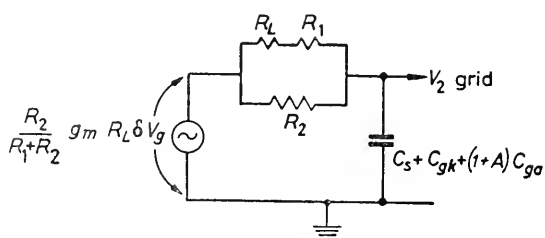


Figure 9.16

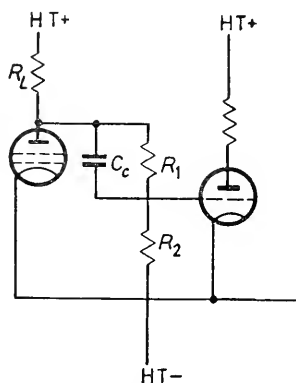


Figure 9.17

Thévenin's theorem to the part to the left, we get *Figure 9.16*. Clearly since—as we have seen— R_1 and R_2 have to be made much larger than R_L the source resistance is higher than in the other systems and the upper cut-off frequency will be correspondingly low.

The solution is to connect a compensating capacitance C_C across R_1 as in *Figure 9.17*. Lumping V_2 input capacitances together as C_{in} , the equivalent

FREQUENCY RESPONSE OF INTERVALVE COUPLINGS

circuit is *Figure 9.18*. Consider the behaviour of the network $R_1 C_C R_2 C_{in}$.

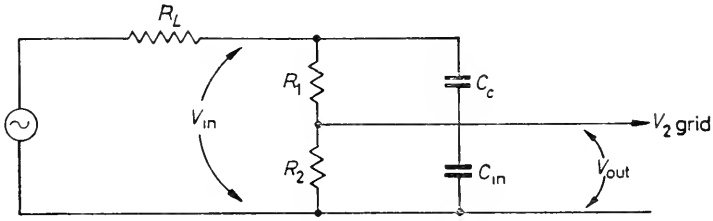


Figure 9.18

If R_L is small in comparison

$$\frac{V_{out}}{V_{in}} = \frac{\frac{R_2/j\omega C_{in}}{R_2 + 1/j\omega C_{in}}}{\frac{R_2/j\omega C_{in}}{R_2 + 1/j\omega C_{in}} + \frac{R_1/j\omega C_c}{R_1 + 1/j\omega C_c}}$$

which simplifies to

$$\frac{V_{out}}{V_{in}} = \frac{R_2}{R_1 + R_2 \left(\frac{j\omega C_{in} R_2 + 1}{j\omega C_c R_1 + 1} \right)}$$

Now let

$$C_c = \frac{C_{in} R_2}{R_1}$$

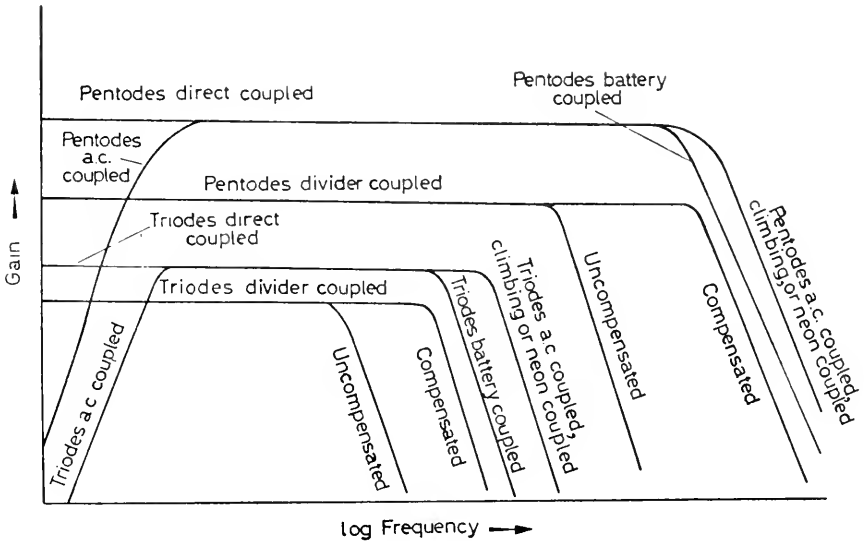


Figure 9.19

If this is substituted into the above, we find the bracketed term reduces to unity. Therefore V_{out}/V_{in} is merely $R_2/(R_2 + R_1)$. It is as if there were no capacitance there and the coupling is not frequency conscious.

What we have done is to arrange that the fraction of the signal transferred from V_1 anode to V_2 grid is independent of frequency. The upper frequency

INTERSTAGE COUPLINGS

behaviour is now limited by the filter made by R_L and C_C in series with C_{in} (Figure 9.18). The turn-over frequency obtained will be comparable with those in other systems. In Figure 9.19 I have summarized the findings of this section. The curves show the form of the frequency response to be expected from the various systems in coupled triodes and coupled pentodes. It is clear that amplifiers are low-pass or band-pass devices. For a complete amplifier comprising a number of couplings in cascade the overall frequency response is the sum of the frequency response of the several couplings.

TRANSIENT RESPONSE OF INTERVALVE COUPLINGS

If a direct coupled amplifier be tested by applying a square wave of voltage at the input, the output does not rise instantly to the final value because of the imperfections in the high frequency response. The output will be of the form of Figure 9.20. The performance of complete amplifiers is some-

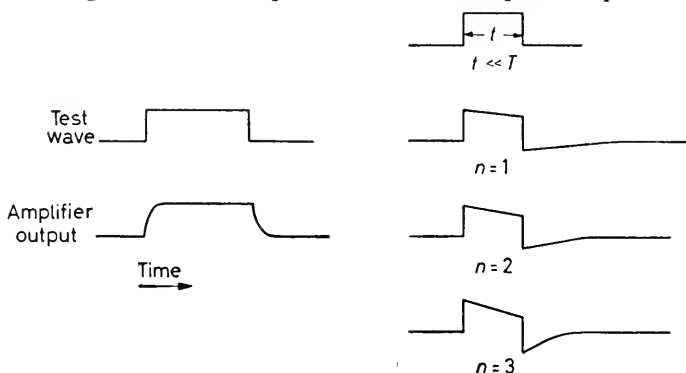


Figure 9.20

Figure 9.21

times given in terms of the time taken by the output voltage to reach, say, 95 per cent of the proper value. This would be called the 'rise time for 95 per cent'. If the amplifier is a.c. coupled there is not only a finite rise time, but also 'sag' due to the non-existence of gain at zero frequency. If all the RC couplings have an identical time-constant T , and if the pulse length $t \ll T$, then if the amplifier contains n RC couplings, the response sags exponentially

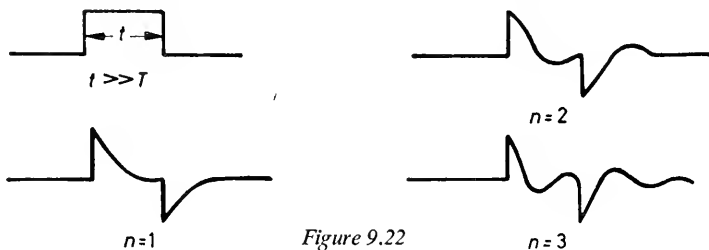


Figure 9.22

with a time constant T if $n = 1$, and approximately exponentially with a time constant T/n if $n > 1$ (Figure 9.21). When $t \gg T$ the response is quite different. The response for $n > 1$ is oscillatory, as is shown in Graph 7 (Figure 9.22).

DIFFICULTIES WITH SINGLE-SIDED AMPLIFIERS

Amplifiers which consist of a simple string of cascaded stages like the 3 stage a.c. coupled amplifier in *Figure 10.1* and the 3 stage direct coupled version in *Figure 10.2* are called single-sided, for a reason which is apparent

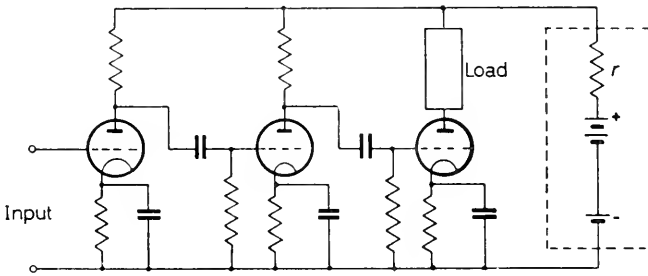


Figure 10.1

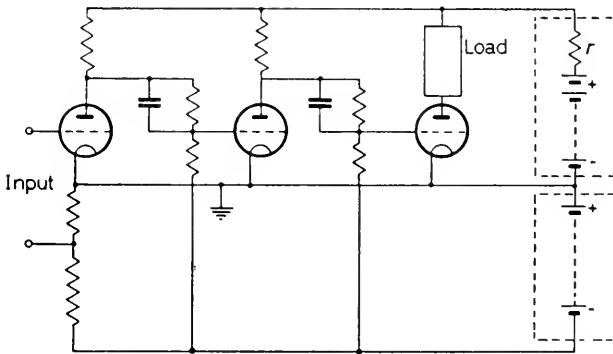


Figure 10.2

later. In operation they present certain difficulties, some of them insurmountable, which render them unsuitable for biological work.

MOTOR BOATING

It is desirable and economical if all the HT power for an amplifier can be supplied from a common source, either battery or 'power pack'. An attempt has been made to do this in *Figures 10.1* and *10.2*, where a common HT battery is shown. It is highly unlikely that these circuits would work—probably they would oscillate violently, no matter what the input signal did. The following give the reasons why.

DIFFICULTIES WITH SINGLE-SIDED AMPLIFIERS

The HT battery inevitably has some internal resistance r . The last valve is a power valve, or a high level amplifier, and passes more standing anode current than the other two. This anode current flows through r and develops a voltage drop across it which will fluctuate as V_3 anode current fluctuates. This varying voltage drop appears as inconstancy of the supposedly fixed HT. A little thought shows that with three triodes (or more) these HT variations are transmitted back to V_1 anode, are then amplified by V_2 and V_3 , and re-appear as variations in V_3 anode current in such a sense as to reinforce themselves. This is an example of undesired positive feedback, and in the early days of radio was called 'motor boating' because of the noise it made when the load is a loudspeaker.

Or again, consider the two valve amplifier of *Figure 10.3*. The first valve

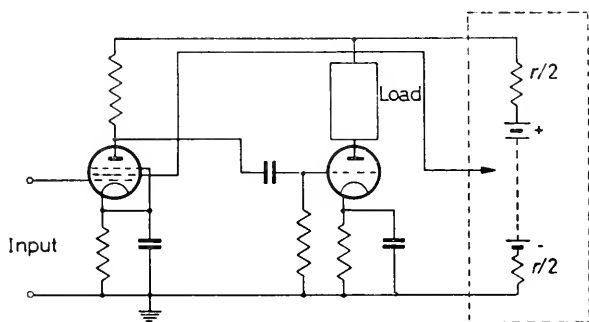


Figure 10.3

is a pentode and the screen is fed from an appropriate tap on the HT battery. The battery resistance is represented as being split into halves at each end. An increase in load current will produce a *reduction* in HT which is fed back to V_1 anode; *but* there is also a reduction of screen potential which will reduce V_1 anode current, and hence the drop in V_1 anode load, so that this tends to *increase* V_1 anode potential. It is likely that the screen effect will be the greater, producing a net positive-ward movement of V_2 grid and further increase in V_2 anode current. This arrangement can therefore also motor boat.

Decoupling

With a.c. coupled amplifiers motor boating can be prevented. Since the amplifier has little gain at low frequencies below the pass-band, if the offending inconstant supply potential is passed through a low-pass filter which also cuts off at a frequency below the pass-band, then motor boating cannot occur. This process is called 'decoupling'. A practical version of *Figure 10.1* is *Figure 10.4*, where the decoupling components are R_a and C_a . Some loss of HT voltage occurs across R_a , and the first stage must be designed with this in mind. It is reasonable to lose between 50 and 100 V across R_a .

Having fixed R_a we find C_a . The higher C_a the more effective the decoupling. The quickest way to find how much is required is by experiment. Similarly, *Figure 10.5* shows a practical version of *Figure 10.3*. R_a is worked out first to give the correct screen voltage and current, then C_a is made large enough

DRIFT

to ensure the amplifier is stable. Do not skimp on decoupling capacitance with battery operated gear. Remember that as the battery gets older the internal resistance will increase and so will the positive feedback. Parsimony here will lead to an amplifier which is stable with new batteries but begins to motor boat while there is still useful life in them.

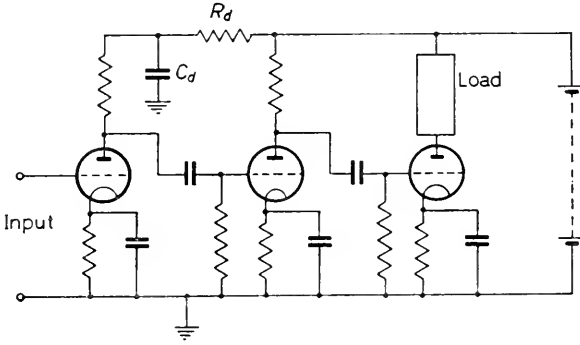


Figure 10.4

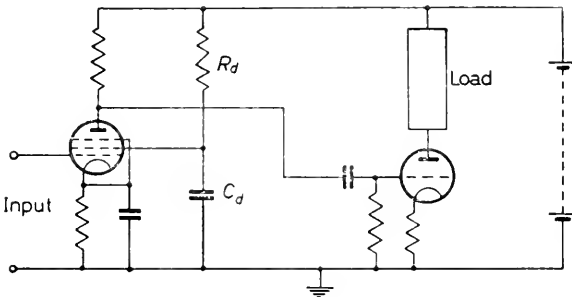


Figure 10.5

With direct coupled amplifiers decoupling does not work because there is no frequency so low that there is no amplification. Furthermore, with long time constant a.c. coupled amplifiers of the kind used in electroencephalography and cardiography, decoupling may be impracticable because capacitances of prohibitive size are required. In these cases it may be possible to do something about motor boating by the widespread use of stabilizer tubes, but it is probably cheaper to begin again with an amplifier of the double-sided type.

DRIFT

Even if the amplifier of *Figure 10.2* could be prevented from motor boating it would be extremely prone to 'drift'. This is the blight with all direct coupled amplifiers and refers to the tendency for the output to alter all the time even when there is no change at the input. It is caused by minor fluctuations in the supposedly constant parameters, such as HT voltage, temperature of valve cathode, and component values generally. Drift is impossible

DIFFICULTIES WITH SINGLE-SIDED AMPLIFIERS

to eradicate altogether, but is likely to be particularly bad in a single-sided amplifier design.

LOAD POLARIZATION

In single-sided power stages directly coupled to a load such as a penwriter, or transformer coupled to a device such as a loudspeaker, there is in both cases a steady component of current which produces a steady magnetic bias in the iron of the penwriter or transformer. It is clear from *Figure 10.6* that this is a bad thing. The curves represent the hysteresis curve of the iron. When there is no bias the working point is at *O* and the input can have the value shown in *Figure 10.6a*. When bias is present the working point is

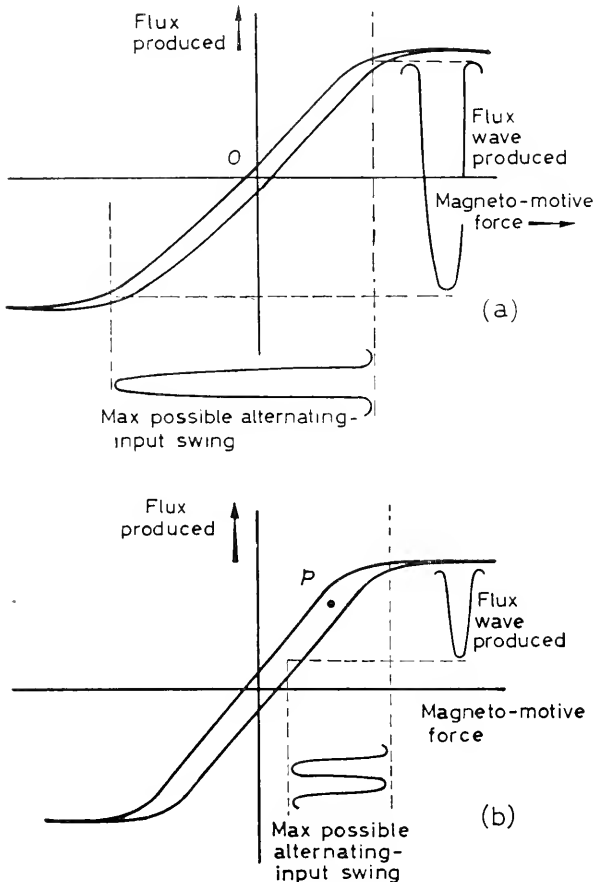


Figure 10.6

some such as *P* in *Figure 10.6b* and saturation occurs if the input exceeds the level shown. The implication is that the power-handling capacity of a given penwriter or transformer is reduced, and that to achieve the same power-handling capacity the cross-section of iron in the magnetic circuit must be

NATURE OF QUANTITY MEASURED

increased in order that the flux density be reduced. The effect is less serious with penwriters because these have a substantial air gap.

NATURE OF QUANTITY MEASURED

Finally, a point important in electrophysiology, the single-sided amplifier measures potentials or potential changes with respect to earth. It cannot measure the difference between two potentials neither of which is earth.

All the difficulties enumerated in this chapter may be overcome or mitigated by the 'double-sided' amplifier, and this will be studied next; before we can do so, however, we shall have to take a superficial look at the vast subject of negative feedback.

NEGATIVE VOLTAGE FEEDBACK AND THE STABILIZED GAIN AMPLIFIER

Suppose we have an amplifier with a gain of A times, and we take a fraction B of the output, feed it back and *subtract* it from the input, then this is called 'negative feedback'. If the input to the amplifier proper is 1 unit, the output is A , the fed-back quantity is AB , so the input to the whole device must have been $1 + AB$. The overall gain is thus $A/(1 + AB)$, and if $AB \gg 1$ the overall gain is approximately $1/B$, independent of A (Figure 11.1).

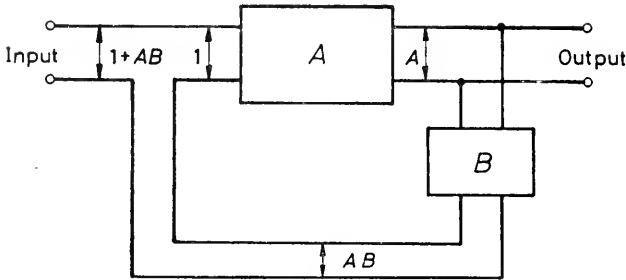


Figure 11.1

Now if B is derived by a potential divider, it is both linear and independent of frequency. Thus an amplifier with negative feedback for which it is possible to say that $AB \gg 1$ ought, at a cost only of reduced overall amplification,

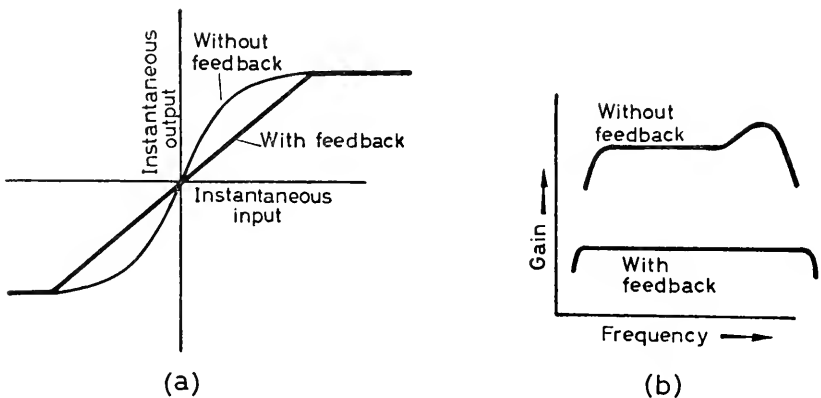


Figure 11.2

to exhibit very level frequency response, constant gain and absence of distortion, and in point of fact this is exactly what does happen with feedback properly applied (Figure 11.2). In a straightforward electronic amplifier

without feedback, the amplitude and frequency distortion may be known and tolerable, but the gain is likely to vary a lot as valves and batteries wear out. If the proposed experiment is designed so that the amplifier is part of a null detector, or if a calibrating input is always available, gain instability does not matter. If not, it does, and gain stability is achieved by negative feedback.

Suppose we have a direct coupled amplifier possessing three similar inter-stage couplings whose upper turn-over frequencies are 10,000 cycles. Suppose the voltage gain is 100,000 times and we try to reduce this to 1,000 by negative feedback. Then $1/B = 1,000$, $B = 0.001$, and $AB = 100$, which is unquestionably much greater than 1, so we ought to be all right. When we try it out, however, we find the whole device oscillates violently at about 17,500 c/s. Why?

In the expression $gain = A/(1 + AB)$, if AB is positive, the feedback is negative and the gain is less than A . If AB is negative but not as negative as -1 , the feedback is positive and the gain is more than A . When $AB = -1$, the gain is infinite, which means that the output can be finite when there is no input, i.e. the device oscillates.

In practice A is complex, as there is both gain and phase shift. Let it be $\alpha e^{j\theta}$, where α is the modulus of the gain and θ the shift phase. Then the gain expression is $(\alpha e^{j\theta})/(1 + \alpha B e^{j\theta})$. In the pass-band of the amplifier θ is small, $e^{j\theta}$ is nearly unity, so that $\alpha/(1 + \alpha B)$ is substantially correct for the gain expression. Outside the pass-band there is appreciable phase shift and $e^{j\theta}$ is important. The gain expression goes to infinity when $\alpha B e^{j\theta} = -1$. As αB is real and positive, $e^{j\theta}$ must be real and negative and the solution for θ must be 180 degrees and for αB , $+1$. αB is called the 'loop gain' of the system, and θ is of course the total amplifier phase shift. *Oscillation will occur if the loop gain exceeds 1 at the frequencies where the total phase shift is 180 degrees.*

In the case under consideration there is one such frequency. At the upper end of the pass-band the intervalve couplings are acting as low-pass filters and introducing phase lag. As there are three of them and they are similar, oscillation occurs when the phase shift of each is 60 degrees, and reference to *Graphs 9 and 10* shows that this happens at $\omega = 1.75/RC$ and that at this frequency $|V_{out}/V_{in}| = 0.5$. As the couplings turn over at 10,000 c/s, oscillation will occur at $1.75 \times 10,000 = 17,500$ c/s, because the loop gain = $100 \times 0.5 \times 0.5 \times 0.5 = 12.5$, which is more than 1.

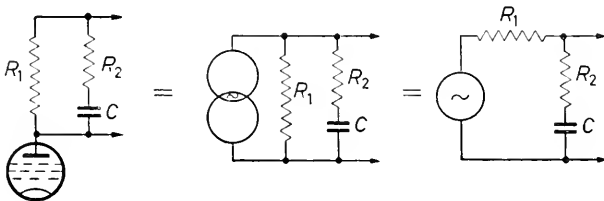


Figure 11.3

To prevent the oscillations we have to reduce the gain at 17,500 c/s whilst leaving it as unaffected as possible in the pass-band, and without introducing any more phase shift. This may be done by modifying the anode load of one of the stages as in *Figure 11.3*, giving us the low-pass circuit which is the

subject of *Graph 12*. C is chosen so that at low frequencies the effect of C and R_2 is negligible, but at 17,500 c/s the reactance of C is small and the two resistances are effectively in parallel. The gain is much reduced. At some intermediate frequency the complex load draws a leading current and thus introduces some phase shift, but this is permissible because the shift in the three coupling circuits is not yet approaching 60 degrees.

Without going into the matter too deeply it may be said that the effect of correcting circuits such as these is to reduce the amplifier pass-band. Therefore to apply successfully feedback to an existing amplifier it is helpful if *the amplifier has a useful gain over a band which is wider than is finally required*. This is the reason for the dictum that efforts to salvage a bad amplifier with negative feedback may be disappointing.

If the amplifier is a.c. coupled there will be another frequency—a very low one—at which oscillations can take place, due to phase advance in the coupling capacitance. To deal with this we connect across the coupling capacitance a correcting network comprising R_2 and C_2 (*Figure 11.4*) producing the high-

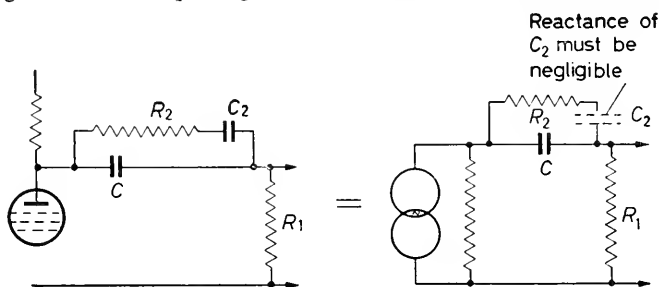


Figure 11.4

pass circuit described in *Graph 12*. At the oscillation frequency R_2 is comparable with the reactance of C , and C_2 must be sufficiently large for its own reactance to be small compared with the resistance of R_2 . The effect is to reduce the phase shift without increasing the loop gain. An alternative approach to the stability problem is to make B deliberately frequency conscious, so that B is substantially real within the pass-band but complex without it and of such a nature as to introduce compensating phase shifts and/or reduce the loop gain at the frequencies where the amplifier would otherwise tend to oscillate.

The conclusion to be drawn from all this is that negative feedback over a number of stages is quite complicated and should not lightly be undertaken. It is, however, perfectly easy to apply round a single stage because the phase shift in the coupling cannot reach 180 degrees, and a similar effect to overall feedback may be had by a number of small feedback loops, as in *Figure 11.5*. Let us see how this may be done.

Suppose we have a triode voltage amplifier, automatically biased, and inadvertently leave the cathode bypass capacitance out (*Figure 11.6*). The load seen by the valve is any resistance through which the anode current passes, and across which it causes voltage fluctuations. Therefore R_K is now as much part of the load as R_L . Now the input to the valve is the grid-cathode voltage and this is evidently $\delta V_{in} - \delta V_{R_K}$, but δV_{R_K} is part of the total

NEGATIVE VOLTAGE FEEDBACK AND THE STABILIZED GAIN AMPLIFIER

output voltage $\delta V_{R_L} + \delta V_{R_K}$. Therefore the input to the valve is the difference between the signal δV_{in} and a fixed fraction of the output, so this is a negative feedback system, the feedback being applied round one stage only. The A

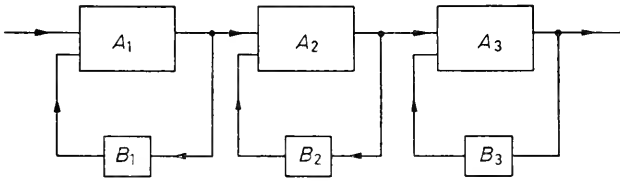


Figure 11.5

is $\{\mu(R_L + R_K)\}/\{r_a + (R_L + R_K)\}$ and the B is $R_K/(R_L + R_K)$. Typically R_L might be 50,000 Ω , $r_a = 10,000 \Omega$, the bias 2 volts, the anode current 2 milliamps and μ equal to 50. From these figures R_K must be 1,000 Ω , so the gain without feedback will be

$$A = \frac{50(50,000 + 1,000)}{(10,000 + 50,000 + 1,000)} \approx 42$$

$$\text{Our } B = \frac{1,000}{50,000 + 1,000} \approx 0.02$$

AB is therefore only 0.84, which is *not* $\gg 1$, and therefore the amount of feedback is small. If in course of time μ falls to half its original value, the gain with feedback falls from $42/(1 + 0.84)$ to $21/(1 + 0.42)$, i.e. from 23 to 15.*

The gain can hardly be said to be stabilized. To improve matters we need more R_K . If we merely increase R_K in Figure 11.6 the valve tries to give

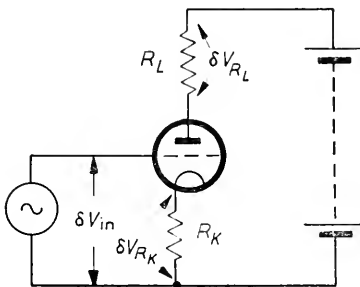


Figure 11.6

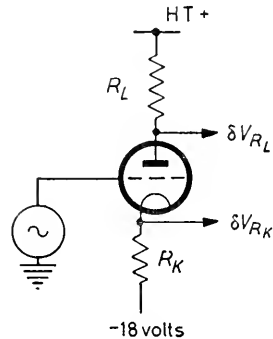


Figure 11.7

itself more bias and the current falls. To increase R_K but preserve the *status quo* in other respects we need a new circuit; we return the bottom of R_K to a fixed negative potential (Figure 11.7). Suppose we make R_K 10,000 Ω .

* Assuming we take the output to be $\delta V_{R_L} + \delta V_{R_K}$. Usually only δV_{R_L} is used, so the gain is effectively lessened by the factor $R_L/(R_L + R_K)$.

NEGATIVE VOLTAGE FEEDBACK AND THE STABILIZED GAIN AMPLIFIER

If the cathode current is still 2 mA, the voltage drop across R_K will be 20 V. If the negative supply is made -18 V the cathode will still be 2 V positive with respect to the grid and the bias is still negative even though it might appear from *Figure 11.7* that it is positive. The gain without feedback is now

$$A = 50 \frac{50,000 + 10,000}{50,000 + 10,000 + 10,000} = 43 \quad B \text{ is } \frac{10,000}{10,000 + 50,000} = 0.167$$

and $AB = 7.2$. If it be allowed that 7.2 is $\gg 1$, then it may be stated straight off that the gain is approximately $1/B = 6$. The exact gain expression is $A/(1 + AB) = 43/(1 + 7.2) = 5.25$. If μ falls to half the new gain is $21.5/(1 + 3.6) = 4.7$. As a result of increasing R_K the change in gain *with feedback* for a halving of valve μ falls from 35 to 10.5 per cent—a desirable improvement.

CONCERTINA PHASE SPLITTER

When $R_K = R_L$, δV_{R_K} and δV_{R_L} are equal and opposite, and the device is called a phase splitter (*Figure 11.8*). When the grid is made positive, for

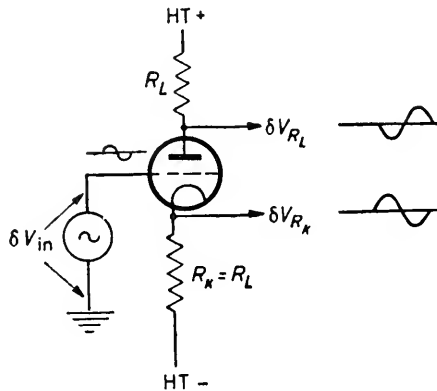


Figure 11.8

example, the valve current increases and so do the drops across R_K and R_L ; therefore the anode goes negative and the cathode positive. As the grid potential alternates the anode and grid may be visualized as going 'in and out' concertina-wise. The gain with feedback is

$$\frac{\delta V_{R_K} + \delta V_{R_L}}{\delta V_{in}} = \frac{2\delta V_{R_L}}{\delta V_{in}}$$

and is $\frac{A}{1 + AB}$ where $B = \frac{1}{2}$ and $A = \frac{\mu 2R_L}{r_a + 2R_L}$.

It is about 2.

The gain, taking the output between either output terminal and earth is $(\delta V_{R_L})/(\delta V_{in})$ and is therefore about unity.

CATHODE FOLLOWER

CATHODE FOLLOWER

When $R_L = 0$ the whole of the load resistance is between the cathode and negative supply so that the feedback fraction is 1 and the gain is 1 approximately or $A/(1 + A)$ exactly (Figure 11.9). If we have the same valve as before and $R_K = 10,000 \Omega$,

$$A \text{ is } \mu \frac{R_K}{r_a + R_K} = 50 \frac{10,000}{10,000 + 10,000} = 25$$

and when μ has fallen to half its proper value $A = 12.5$. The gain alters from $25/(1 + 25)$ to $12.5/(1 + 12.5)$ or from 0.945 to 0.93, a change of only

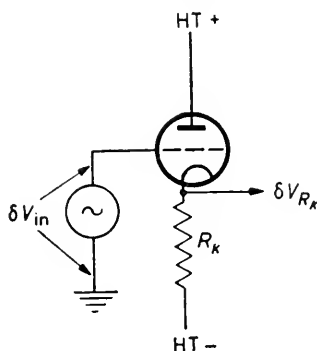


Figure 11.9

1.5 per cent. Undoubtedly a very high degree of gain stabilization; it may well be asked, however, what is the point of taking all this trouble with a device whose gain is less than 1, which in fact attenuates slightly. The answer is that the cathode follower exhibits *par excellence* two valuable results of applying negative feedback which we have not so far mentioned: high input impedance and low output impedance.

Input impedance

This comprises two elements in parallel, the input capacitance and the input resistance.

Input resistance—We have to digress for a moment to properties of valves in general. Up till now we have assumed that the input resistance of a valve is infinite. In most valve applications this is to all intents and purposes true, for the input resistance is certainly sufficiently higher than the source or generator resistance to be neglected. However, in a few applications such as micro-electrode recording in electrophysiology or pH measurement, the generator (micro-electrode or pH probe) resistance is itself extremely high, and the valve input resistance becomes important.

In point of fact a small but significant grid current, which may lie in the range 10^{-6} to 10^{-12} amps with ordinary radio receiving type valves, flows either into or out of the valve, depending on the bias used. The current is due to the net effect produced by a number of causes—electrons hitting the

NEGATIVE VOLTAGE FEEDBACK AND THE STABILIZED GAIN AMPLIFIER

grid, ions hitting the grid, leakage of current through surface moisture on the valve envelope, and the whole lot combine to produce a curve having the general shape of *Figure 11.10*. The slope of the curve at any particular

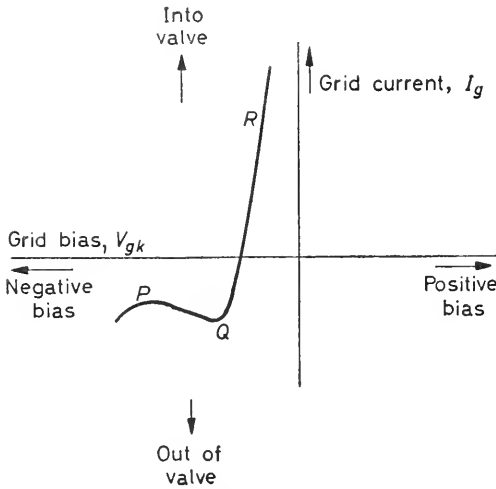


Figure 11.10

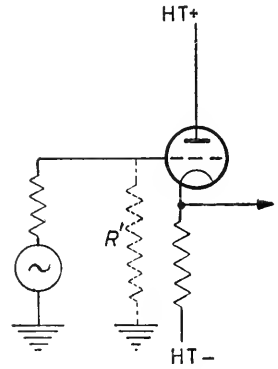


Figure 11.11

value of bias gives the value of the input resistance. The input resistance is positive between *Q* and *R*, infinite at *Q* and negative between *P* and *Q*.

What does negative resistance mean? The power absorbed by a positive resistance is V^2/R , so the power absorbed by a negative resistance is $-V^2/R$, that is, it does not absorb power, it generates it. We cannot have spurious generators associated with the input circuits of measuring devices, and the net input resistance must somehow be made positive. The spurious generator is most powerful when the negative resistance is *lowest*. If this value is $-R_{i\text{-min}}$, then if we connect between grid and cathode an ordinary resistance R_g , the resistance of the parallel combination is

$$\frac{-R_i R_g}{R_g - R_i}$$

and this is positive if $R_g < R_i$.

This is the explanation for the maximum value of grid resistance permissible in amplifying stages being stated by the valve manufacturers. It means that the valve can be operated over the bias range corresponding to the *PQ* section of the grid current characteristic without instability. Over the *QR* range there is no difficulty anyway, as the input resistance is positive.

Now let us see how this affects the cathode follower. Suppose we are working in the region where the input resistance is positive. If we make the grid positive, say, an amount δV_g , then in an ordinary amplifying stage the change in grid current is δI_g and the input resistance is $R_i = \delta V_g / \delta I_g$. In a cathode follower having a gain of, say, 0.95, when we apply δV_g we also produce a change of cathode potential $\delta V_k = 0.95 \delta V_g$, and so the change of grid current is only $0.05 \delta I_g$. The apparent input resistance is thus $\delta V_g / (0.05 \delta I_g)$, 20 times higher than when the valve is used as an amplifier. By a similar

CATHODE FOLLOWER

argument, if we are working where the input resistance is negative, then it is 20 times higher in the cathode follower configuration than as an amplifier, and the amount of external positive resistance required to ensure stability can be 20 times higher. This much larger external resistance can often be the source resistance itself, and there is then no need to provide a special resistor (R' in *Figure 11.11*) to ensure stability. Thus, whatever the bias used, it is true to say that the cathode follower has a high input resistance.

A.c. coupled cathode follower

When cathode followers appear in a.c. coupled amplifiers, the arrangement is slightly different (*Figure 11.12*). Automatic steady bias is developed

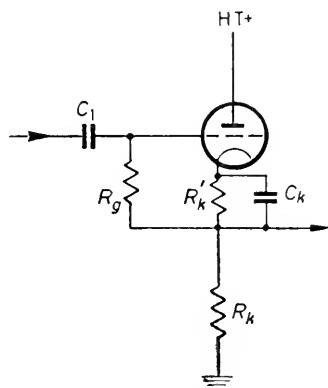


Figure 11.12

across $R'_k C_k$. R_g , because it is returned to a point which moves with the cathode and not to earth, is the ordinary upper limit of grid resistance allowed by the valve makers. For a gain of 0.95 as before, however, the effective value of R_g as seen from the source is $20 R_g$, so the input resistance is very high. Furthermore, $20 R_g$ is the resistance value relevant to the calculation of required coupling CR product, so C_1 can be 1/20 of the value it would have to have if the valve were an amplifier.

Input capacitance

With the cathode follower circuit the capacitance between grid and anode remains unaffected at C_{ga} (there is no Miller effect because the anode potential is fixed) but the effective capacitance between grid and cathode is much reduced (*Figure 11.13*). If there is a change of grid potential δV_g and the cathode follower gain is 0.95, $\delta V_k = 0.95 \delta V_g$ and $\delta V_{gk} = 0.05 \delta V_g$. The displaced charge is $0.05 \delta V_g C_{gk}$ so the apparent grid-cathode capacitance is

$$\frac{\text{Displaced charge}}{\text{Change of grid voltage}} = \frac{0.05 \delta V_g C_{gk}}{\delta V_g} = 0.05 C_{gk}.$$

The grid-cathode capacitance is therefore reduced 20 times.

With a pentode cathode follower the gain without feedback is higher than with a triode, so the gain with feedback is more nearly unity and the improvement in input resistance and grid-cathode reactance even better. The gain

NEGATIVE VOLTAGE FEEDBACK AND THE STABILIZED GAIN AMPLIFIER

without feedback can easily be 100, so that cathode follower gain will be $100/(1 + 100) = 0.99$. The factor by which the input resistance and grid-cathode reactance is improved is now approximately 100. Further, since the pentode has a very low C_{ga} , there is also a big increase in the grid-anode reactance.

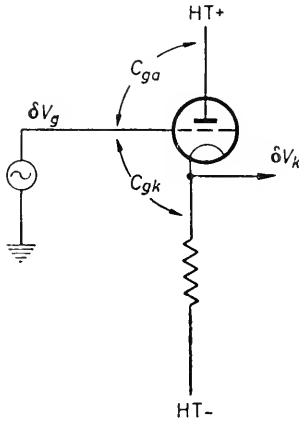


Figure 11.13

There is a minor complication. Pentode action depends on the potential difference between cathode and screen remaining constant. Cathode-follower action requires that the load current be equal to the anode current, or at least different from it only by a constant amount (Figure 11.14). This means

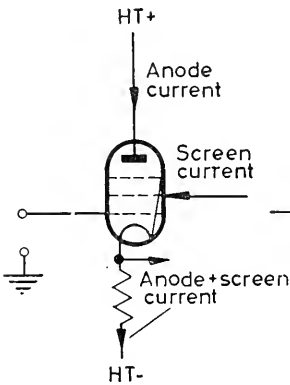


Figure 11.14

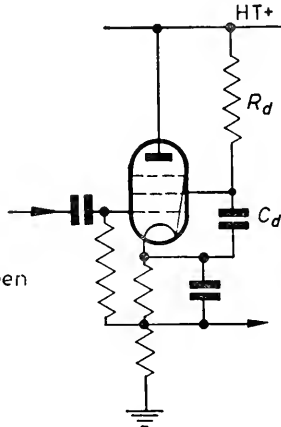


Figure 11.15

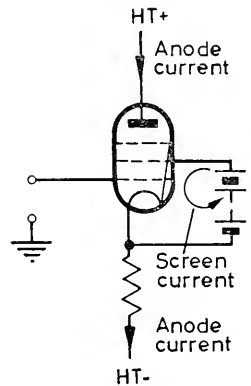


Figure 11.16

that the screen current must be held constant, which again requires that the screen-cathode voltage be somehow held steady. In the a.c. coupled pentode cathode follower this may be achieved by returning the screen decoupling capacitance to the cathode and not to earth (Figure 11.15). In the direct coupled pentode cathode follower the screen is most conveniently supplied

CATHODE FOLLOWER

from a special battery which is returned to the cathode. This method has the advantage that the screen current is kept out of the load (*Figure 11.16*).

Advantage of high input impedance

If our source of input voltage has a high internal resistance r , and the valve into which it works has input resistance R and input capacitance C (*Figure 11.17*), then there occurs:

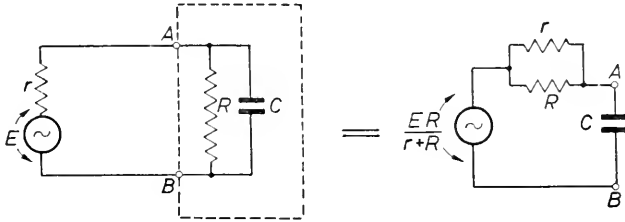


Figure 11.17

(1) A loss of voltage delivered to the input terminals AB , due to the potential divider action of R and r ; the fraction of the e.m.f. transferred is $R/(r + R)$.

(2) An additional loss at high frequencies due to the shunting effect of C . We have a simple low-pass filter turning over at

$$\omega_c = \frac{1}{C \left(\frac{rR}{R + r} \right)}$$

To get ω_c high we need r , R and C low. We have no power to alter r ; R must be $\gg r$ for reason (1), so C must be small.

With the cathode follower C is small and R is large, so it forms the ideal stage to follow generators of very high internal resistance. A conventional amplifier may then follow the cathode follower.

Output impedance

The output impedance of a device is its internal resistance regarded as a generator, and may be inferred by seeing what happens to the output voltage when some kind of external load is connected across it.

In the case of the cathode follower, suppose as the result of the connection of a load R_L (*Figure 11.18*) the output voltage falls δV_k . The grid potential is unaffected, so $\delta V_{gk} = \delta V_k$ and an extra current flows in the valve $\mu \delta V_k / (R + r_a)$ where $R = R_k$ in parallel with R_L . R_L is invariably much smaller than R_k , so the extra current is approximately $\mu \delta V_k / (R_L + r_a)$ and nearly all flows in the external load.

The output impedance is

$$\frac{\text{Change of output voltage}}{\text{Change of output current}} \approx \frac{\delta V_k}{\frac{\mu \delta V_k}{R_L + r_a}} = \frac{R_L + r_a}{\mu} = \frac{R_L}{\mu} + \frac{1}{g_m}$$

NEGATIVE VOLTAGE FEEDBACK AND THE STABILIZED GAIN AMPLIFIER

If the valve is a pentode $R_L \ll r_a$ and the output impedance is given with sufficient accuracy merely by $1/g_m$. For both triode and pentode it is real and positive, meaning that the output impedance is resistive. g_m is typically 5 mA/V, giving an output resistance of only 200 Ω . This is very low compared with the output resistance of an amplifying stage, which is for a pentode substantially equal to the load, say 100,000 Ω (Figure 11.19), and

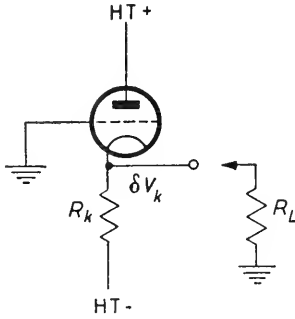


Figure 11.18

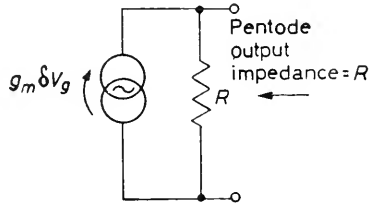


Figure 11.19

for a triode $(R \cdot r_a)/(R + r_a)$, where R is typically 30,000 Ω and r_a 10,000 Ω , giving an output resistance of 7,500 Ω (Figure 11.20). The cathode follower is the nearest approach in electronics to a constant-voltage generator for signals.

Advantage of low output impedance—Many of the resistance-capacitance networks we mentioned in Chapter 4, such as the twin T , were analysed on the basis of their being driven from a constant-voltage generator and feeding a load of infinite resistance. The nearest we can get to this state of affairs in practice is to work the device between two cathode followers (Figure 11.21).

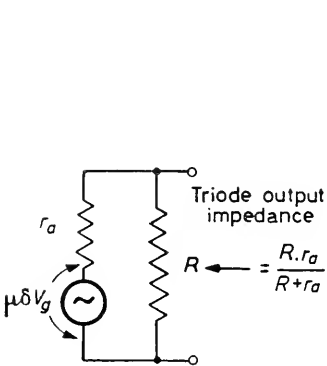


Figure 11.20

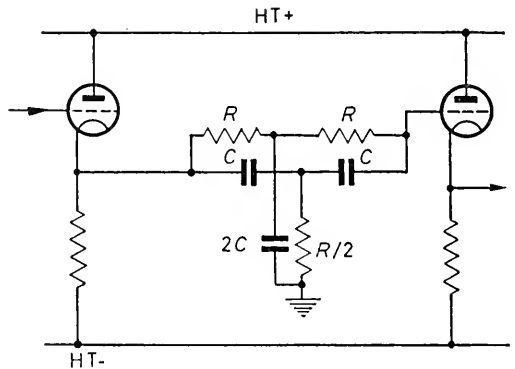


Figure 11.21

The input to the first cathode follower can then be derived from some circuit having rather indifferent properties for constant-voltage generation. Similarly the output of the second may be passed on to something having quite a low resistance. Again in Figure 11.22, suppose device A is feeding device B via a rather long piece of screened cable having a capacitance between

CATHODE FOLLOWER

the central conductor and the outside braiding of $1,000 \mu\mu\text{F}$. If the output of A is taken from the anode of a pentode with a load of $100,000 \Omega$, then this source resistance with the cable capacitance form a low-pass filter which will

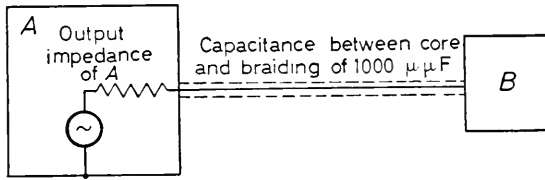


Figure 11.22

turn over at $\omega_c = 1/(10^5 \times 10^{-9}) = 10^4$ radians/sec. This is a frequency of only $10^4/2\pi$ c/s, so the high frequency performance of the apparatus is liable to be disappointing. Suppose the pentode output is passed to the cable via a cathode follower of output resistance 200Ω , then the cut-off frequency will be $1/(2 \times 10^2 \times 10^{-9}) = 5 \times 10^6$ radians/sec. This is quite a different story, and amply high for biological work.

DOUBLE-SIDED AMPLIFIERS

PUSH-PULL VOLTAGE AMPLIFIER

In *Figure 12.1* two similar voltage amplifying stages are shown back-to-back. If instead of two separate signal sources it can be arranged that the same generator supplies equal and opposite inputs to the two valves, from a phase splitter, or in some such manner as shown in *Figure 12.2*, then the valves are

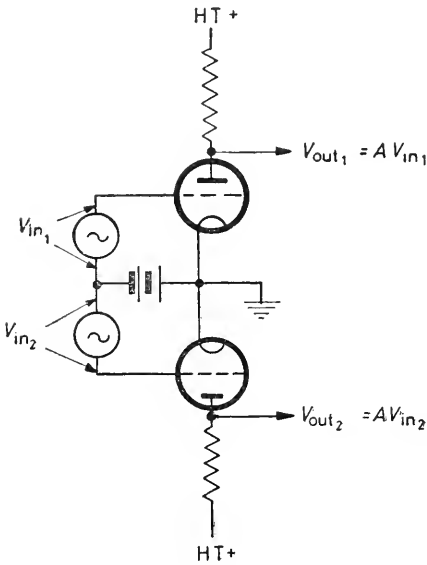


Figure 12.1

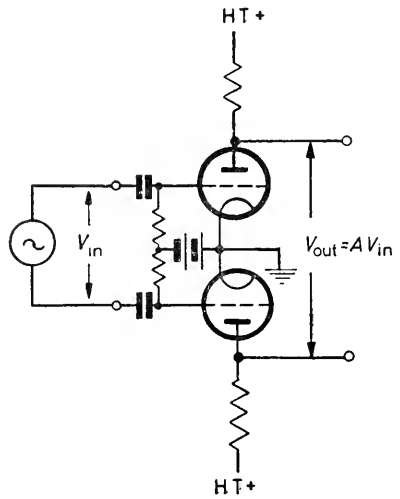


Figure 12.2

said to operate in push-pull. V_{out} is taken between the anodes. The gain is the same as for one valve only, which might at first seem rather wasteful of valves. In fact some very useful results follow.

(1) *Constancy of total anode current*—When the anode current of one side increases that of the other side decreases by a similar amount; the total anode current is therefore constant. If the stage is a high level one there is therefore no tendency for fluctuations to be impressed on the HT voltage, with consequent motor boating.

(2) *Comparative immunity to supply voltage variations*—Since the output is taken from between the two anodes, any change of heater voltage of HT which affects both sides in the same manner does not, to a first order, alter the difference between the anode potentials.

(3) *Grid bias easily applied*—Automatic grid bias is easily applied for direct coupled amplifiers (where applicable) as well as a.c. coupled. Since the total

PUSH-PULL POWER STAGE

anode current is constant the bias developed does not fluctuate and no cathode bypass capacitance is required. The resistance required is half that for a single-sided amplifier because it passes twice the standing current (*Figure 12.3*).

(4) *Screen potential easily applied*—If the valves are pentodes, the screen potential may be derived from the HT by a simple resistance, no decoupling capacitance (a.c. coupled amplifiers) or stabilizing tube (direct coupled)

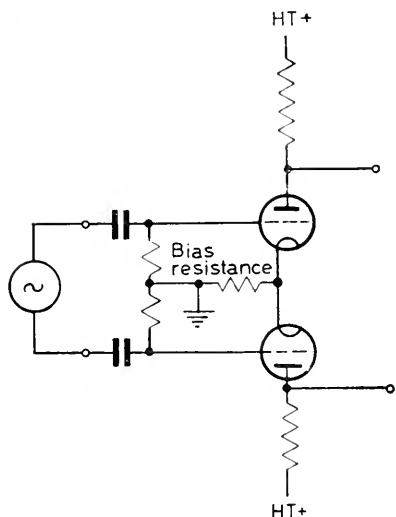


Figure 12.3

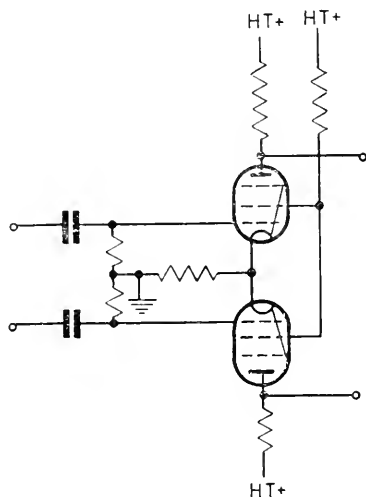


Figure 12.4

amplifiers) being required. This is because the screen currents of the two valves also alter equally in opposite directions. The total screen current is unaffected and the common screen potential may be supplied as in *Figure 12.4*.

The appropriate stage to follow a push-pull amplifier is another push-pull stage, either a further voltage amplifier or a push-pull power stage. Coupling is achieved by any appropriate method from the selection available for single-sided amplifiers.

PUSH-PULL POWER STAGE

This is shown in the a.c. coupled 'class A' (see below) version (*Figure 12.5*) and a direct (neon) coupled version (*Figure 12.6*). The high tension supply is fed to a centre-tap on the transformer primary winding or load respectively. Push-pull power stages have all the advantages already listed for push-pull voltage amplifiers, and three more in addition.

(1) *Absence of load polarization*—We have seen that if the load is a magnetic device, or is coupled to the valves via a magnetic device (transformer), then an undesirable reduction of the handling capacity of the load or transformer resulted from the steady magnetization produced by the mean anode current. In the push-pull output stage the mean anode currents flow in opposite directions and the net standing magnetization is zero.

(2) *Cancellation of certain forms of distortion*—Valve anode characteristic curves—even in the working region—are neither perfectly straight nor evenly

DOUBLE-SIDED AMPLIFIERS

spaced, so the output is not a perfect copy of the input; the problem is more serious with high level and power amplifiers, where the grid is swung over a wide range of voltage. In particular, triode characteristic curves tend to

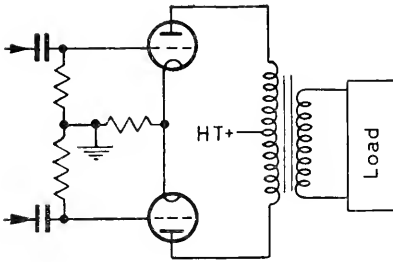


Figure 12.5

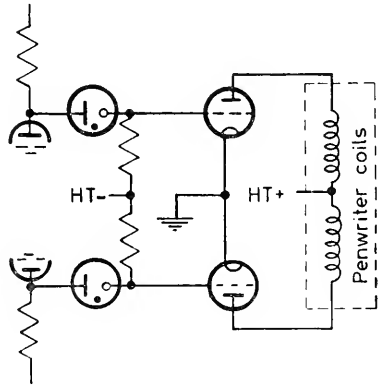


Figure 12.6

bunch in the bottom right-hand corner of the working region (Figure 12.7) and pentode curves at the bottom right and top left (Figure 12.8). The effect of this is that if the valve be fed with a sine wave input, the triode output

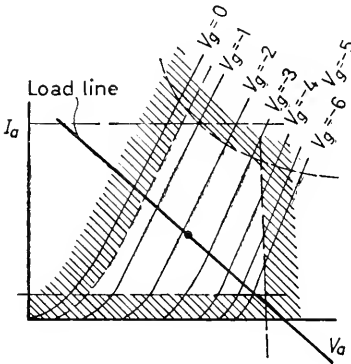


Figure 12.7

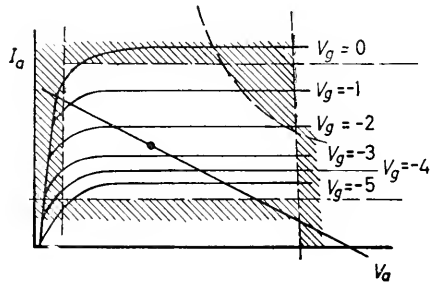


Figure 12.8

tends to be 'squashed' on the positive peaks and the pentode output to be squashed on both positive and negative peaks (Figures 12.9 and 12.10). A waveform of the Figure 12.9 type is produced by adding to a pure sine wave a



Figure 12.9

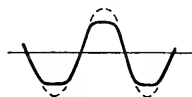


Figure 12.10

small proportion of its even harmonics, notably the second (Figure 12.11). Similarly, a waveform of the Figure 12.10 type is produced by adding to a pure sine wave a small proportion of its odd harmonics, notably the third

PUSH-PULL POWER STAGE

(Figure 12.12). This is the reason for the dictum that second harmonic distortion is associated with triodes and third harmonic with pentodes.

With the push-pull output stage second harmonic distortion (and higher

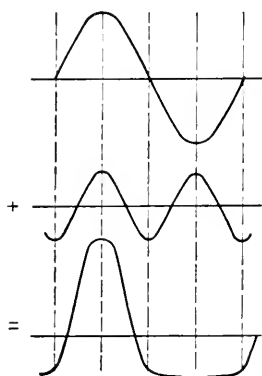


Figure 12.11

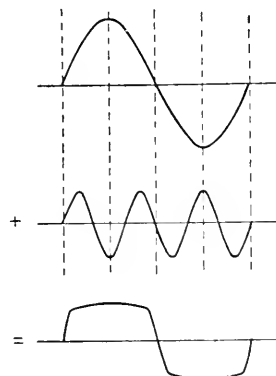


Figure 12.12

even harmonics) cancels out, as indicated in Figure 12.13. Since the inputs are in anti-phase, when one valve is working in the distortion region the other is not, the contributions towards the output of the two valves add, as shown, to produce an undistorted total output. Unfortunately third (and higher odd)

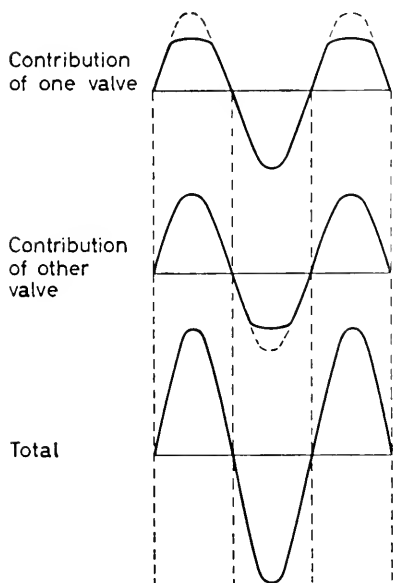


Figure 12.13

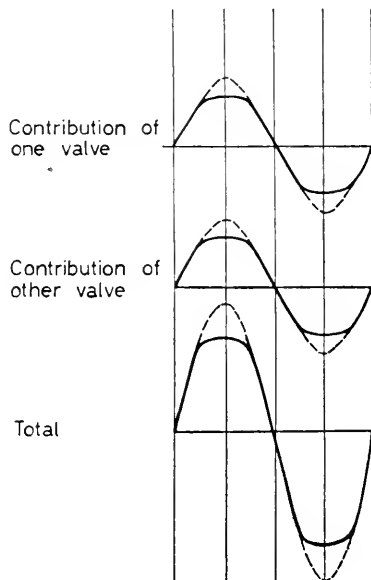


Figure 12.14

harmonic distortion does not, as shown in Figure 12.14. It is thus clear that the best type of output stage for all-round performance ought to comprise triodes in push-pull, and at one time this view was frequently held. In modern

DOUBLE-SIDED AMPLIFIERS

designs, however, pentodes are used because of the advantages they possess—enumerated earlier—and the output waveform distortion is controlled by the application of suitable negative feedback.

(3) *Possibility of securing high efficiency—the class B stage*—The power amplifiers which we have discussed so far have been biased to the middle of the working region. Operation in this mode is called class A. In single-sided circuits class A working is essential but with the push-pull arrangement it is not. With class A push-pull output stages there is a considerable standing anode current which flows whether there is any signal input or not and consequently represents a waste of power.

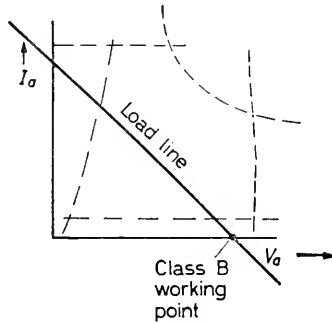


Figure 12.15

In class B working the valves are biased to anode current cut-off, so that the working point of each valve is near the bottom right-hand corner of the working region (Figure 12.15). The standing anode current in the absence of a signal at the input is therefore zero. On the arrival of a signal, each valve

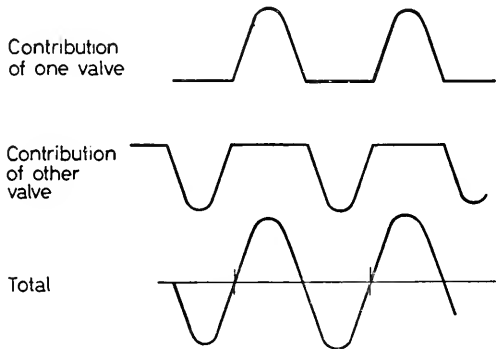


Figure 12.16

conducts alternately—passing a large anode current when the grid is swing positive and remaining cut-off when the grid goes negative (Figure 12.16). High tension power is consumed by the stage only when required, and it is therefore extremely efficient.

DIFFERENTIAL VOLTAGE AMPLIFIER

Class B amplifiers unfortunately suffer from having their mean working point outside what would generally be regarded as the working region, where there is curvature at the foot of the characteristic curves. In consequence there is distortion of the output waveform near the base line. This distortion is called 'crossover distortion' and takes the form of flattening of the bases of the half-sine wave loops due to the cramping of the characteristic curves. A more realistic total output waveform is that of *Figure 12.17*. Crossover

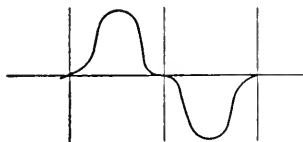


Figure 12.17

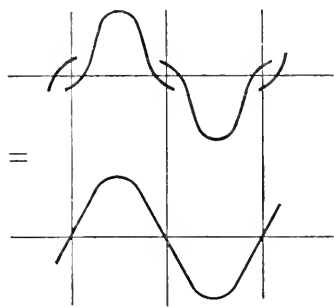


Figure 12.18

distortion is overcome by reducing the bias slightly so that the mean working point moves into the working region: the half-loops then have the form of *Figure 12.18* and the distortion cancels. Some standing anode current now flows and so the stage is slightly less efficient. This mode of working is called class AB. Alternatively, if efficiency is vital (as in miniature battery operated gear) it may be better to reduce crossover distortion with negative feedback.

DIFFERENTIAL VOLTAGE AMPLIFIER

If we take a direct coupled push-pull voltage amplifier and enormously increase the common cathode resistance beyond the sort of value that would be used for automatic biasing, and if we maintain the anode currents by returning this large resistance to a negative supply, a very useful device known as a 'long-tailed pair' results (*Figure 12.19*).

We define an 'in-phase input' as one which moves the grids an equal amount in the same direction (i.e. both positive or both negative) and a 'balanced anti-phase input' as one which moves them an equal amount in opposite directions.

When a balanced anti-phase input is applied to a long-tailed pair the combined cathode current remains the same and no change in voltage drop occurs along the cathode resistance. So far as the signal is concerned the circuit behaves as if R_k were not there and amplification of the input occurs. If the signals applied to the grids are $+\delta V_g$ and $-\delta V_g$, then the potential changes at the anodes are plus and minus $(\delta V_g \mu R_L)/(R_L + r_a)$ (triodes) and $\delta V_g g_m R_L$ (pentodes) in the usual way.

When an in-phase signal is applied the situation is quite different. Both valves change their anode current similarly and the two halves of the circuit are effectively in parallel: an equivalent circuit is given in *Figure 12.20*.

DOUBLE-SIDED AMPLIFIERS

R_k becomes part of the load and introduces negative-feedback. We have already seen that the gain of this circuit when defined as $\delta V_a/\delta V_g$ is only

$$\frac{\frac{R_L}{2}}{\frac{R_L}{2} + R_k} \cdot \frac{A}{1 + AB}$$

where A is $\frac{\mu\left(\frac{R_L}{2} + R_k\right)}{\frac{R_L}{2} + R_k + \frac{r_a}{2}}$ (triodes) or $2g_m\left(\frac{R_L}{2} + R_k\right)$ (pentodes)

and B is $\frac{R_k}{\frac{R_L}{2} + R_k}$

If $AB \gg 1$, as it usually is, this reduces to $R_L/2R_k$. Thus if we make $R_k = R_L$, the changes of potential produced at the anodes by an in-phase input δV_g are in the same direction and equal to $\frac{1}{2} \delta V_g$, and the more R_k is, the less the

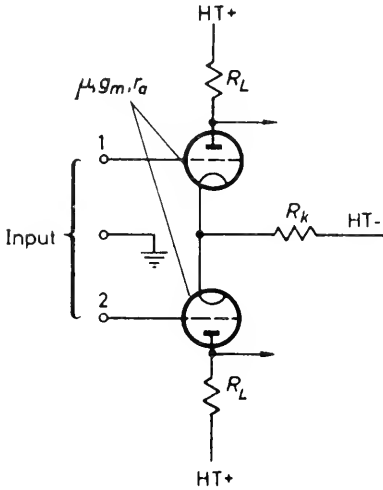


Figure 12.19

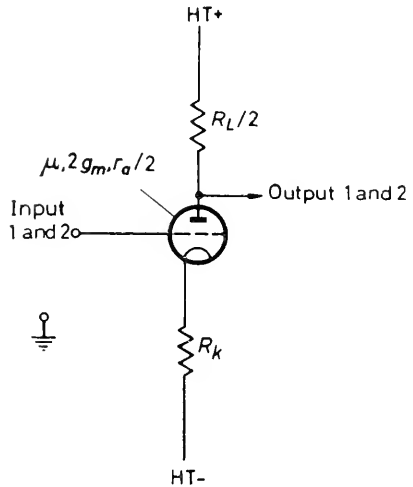


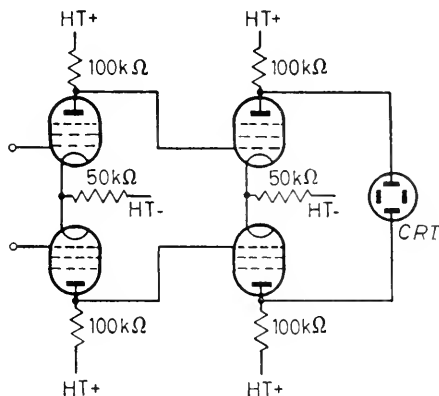
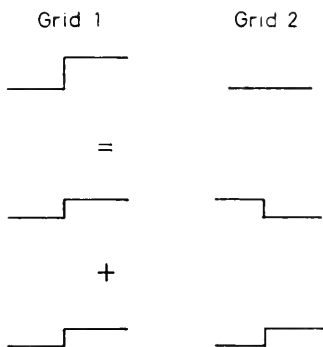
Figure 12.20

anode potential change produced. This is a most useful fact, for it means that we have a *differential amplifier*, i.e. one which responds to inputs which affect the difference of potential between the input grids, but is practically unaffected by their absolute potential. Since the small output which does appear from an in-phase signal is also in-phase, if another stage of amplification by a long-tailed pair follows, this additional stage will further discriminate against the original unwanted in-phase signal.

Further, suppose an input δV_g be applied to *only one grid* of a long-tailed pair. From Figure 12.21 it is clear that this is the same as if we had applied

DIFFERENTIAL VOLTAGE AMPLIFIER

an anti-phase signal $\delta V_g/2$ to both grids, quickly followed by an in-phase signal $\delta V_g/2$. The former causes the anodes to move in opposite directions and the latter has almost no effect. If the anodes move equally in opposite



directions we have all the benefits of push-pull voltage amplification (no decoupling required etc.) without the need for a special phase-splitter to drive the stage. It makes no difference in what ratio the input is divided between the grids, the effect at the output is substantially the same. In other words, the stage responds similarly to an anti-phase input whether the input is balanced or not.

The *discrimination ratio* or rejection ratio of a complete differential amplifier is the ratio of the deflections produced at the indicating device by a unit anti-phase and a unit in-phase signal. Its magnitude depends on the way the indicating device is fed; since the anodes of the last differential stage in the amplifier move equally in opposite directions for any anti-phase signal it is possible, convenient but theoretically undesirable to take the output from between *one* anode only and earth. To take an example, suppose the indicating device is a cathode ray tube of the classical type having a pair of Y deflector plates, and suppose we have two similar differential voltage-amplifying stages employing pentodes having $g_m = 1.5$ mA/volt, anode loads of 100 kΩ and a cathode resistance in each stage of 50 kΩ (as indicated by the incomplete circuit in Figure 12.22). Then the two deflector plates are connected to the final stage anodes and the deflection produced is proportional to the difference of final anode potential.

The stage gain for an anti-phase signal is $g_m R = 150$, so the anti-phase gain of the whole amplifier is 22,500. The in-phase stage gain is $R_L/(2R_k) = 1$. Since for the in-phase signal the final anodes move in the same direction, and since the C.R.T. spot is deflected by *difference* of deflector plate potential, *no* deflection is produced by the in-phase signal and the discrimination ratio is theoretically infinite.

Now suppose we change the cathode ray tube for one of the Cossor double-beam type. In the latter tube each spot is deflected by varying the potential of *one* deflector plate with respect to a beam splitter plate whose

DOUBLE-SIDED AMPLIFIERS

potential is fixed, that is, the tube is suited only for 'asymmetric' deflecting drive. In this case we have no option but to take the output from one anode only (Figure 12.23). The anti-phase gain is now only 11,250, because half the output is wasted at p , but equally the in-phase gain is only $\frac{1}{2}$ and the rejection ratio is $11,250/\frac{1}{2} = 22,500$.

In point of fact it will not be as high as this, and the rejection for the symmetrical case will certainly not be infinity; we have assumed hitherto that

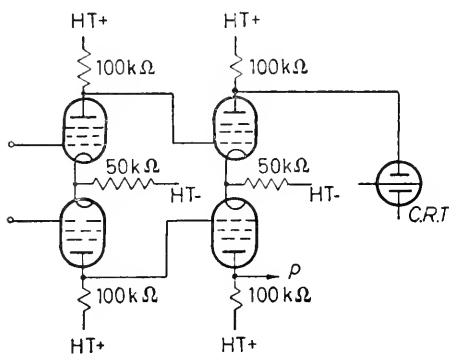


Figure 12.23

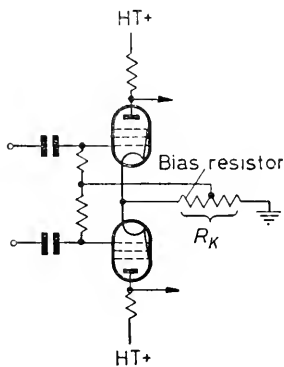


Figure 12.24

the two sides of the amplifier are perfectly similar in all respects. In a practical circuit with real-life components this cannot be the case. What happens is that an in-phase input produces at the first stage anodes potential changes which are in the same direction but not quite equal—that is, an in-phase output with a small anti-phase component. This anti-phase component will suffer amplification by the second stage, so that the practical rejection ratio is always inferior to the theoretical.

The situation can be mitigated by introducing special gain-balancing arrangements. The principal offenders in causing inequality between the two sides are the valves, whose g_m values are subject to rather wide variations both in manufacture and in aging. The object of gain balancing networks is to correct for differences in the performance of the two valves comprising each stage; hence theoretically each stage requires a balancing adjustment. In practice, amplifiers must not have too many knobs; it is clear that since lack of gain balance in the first stage—when handling an in-phase input—leads to amplification of the concomitant anti-phase component by all the succeeding stages, the first stage is the one which requires balancing most, and in practice the provision of first-stage gain-balancing only is usual.

In the a.c. coupled differential voltage amplifier no HT negative supply with respect to earth is necessary, and the stage has the form of Figure 12.24.

MAKING R_K A PENTODE

A good ratio of anti-phase to in-phase gain with a differential stage depends on having R_K large. Unfortunately, if R_K is an ordinary resistance and is to be large, the negative supply must be very negative indeed if the valve currents

GAIN AND BALANCE CONTROLS

are to be maintained. This is inconvenient, but can fortunately be overcome by making R_K a pentode valve (Figure 12.25), for the pentode has a very high incremental resistance (r_a) but quite a low d.c. resistance (V_a/I_a). The

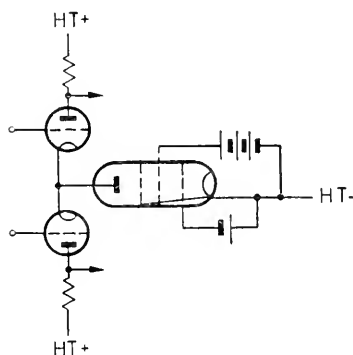


Figure 12.25

effective value of R_K will be the r_a of the valve—possibly 1 megohm—yet this may be achieved with an anode-cathode potential, and hence negative supply of perhaps only 20 V. If the pentode anode current is 1 milliamp, the performance achieved is the same as if a 1 megohm resistance had been used, and a 1,000 V negative supply.

This arrangement is undoubtedly capable of yielding extremely high discrimination ratios but it is necessary to issue a warning. To begin with, there is no point in having a splendid theoretically possible rejection ratio if the stage balancing arrangements are not elaborate enough to secure it. Again, it will be shown in the chapter on interference that the 'equal' in-phase signals cannot be relied upon to be all that equal. If this is so there is no point in striving after enormous rejection ratios. A rejection of 10,000 is usually considered ample and this is easily attained without employing prodigiously negative supplies. In-so-far as an extra valve and its attendant bias, screen and heater supplies are required, the author is of the opinion that, for general purposes at least, the use of a pentode for the cathode resistance is unjustified.

GAIN AND BALANCE CONTROLS FOR DIFFERENTIAL VOLTAGE AMPLIFIERS

Gain

The purpose of gain control is to maintain the output of an amplifier at a convenient level despite wide variations in the magnitude of the input. In biological amplifiers it is achieved either by varying the gain of one of the stages of amplification by negative feedback or by incorporating a potentiometer in one of the interstage couplings, or both.

The point in the amplifying chain at which control of gain takes place requires careful choice. If the point is in the early, low-level stages of the amplifier, then when the gain control is turned down there is no reduction in the noise generated by later stages; the signal-to-noise ratio is poor at low-gain settings. If the control point is in the late, high-level stages, then valves

DOUBLE-SIDED AMPLIFIERS

in earlier stages may be driven outside their working regions—with attendant distortion—by large but not excessive input signals.

If the amount of amplification present is not great, and noise presents no problem, the gain control may confidently be placed at an early stage, possibly even directly after the input terminals (*Figure 12.26*). For high-

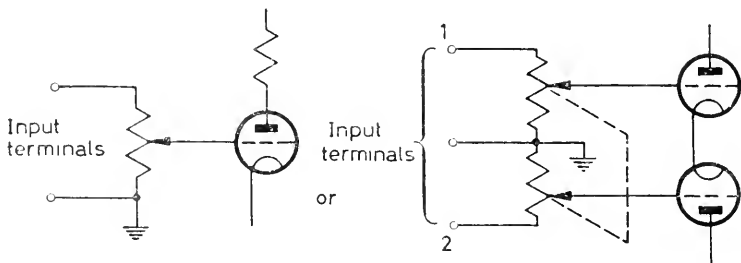


Figure 12.26

gain electrophysiological and similar amplifiers the author favours controlling the second amplifying stage, either the second stage input potentiometrically or the second stage gain by feedback.

A.c. amplifiers—*Figure 12.27* shows in outline a potentiometric continuously

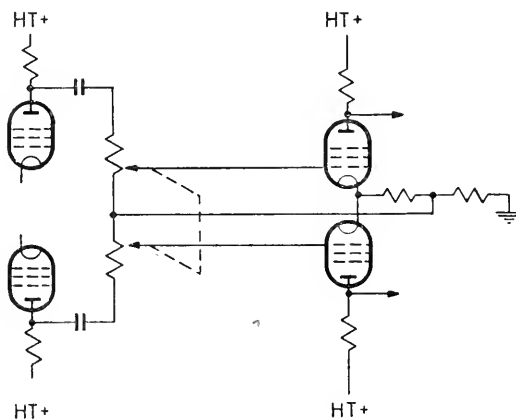


Figure 12.27

variable gain control for an a.c. coupled amplifier which ought theoretically to be satisfactory. In practice it is not, because the ordinary double potentiometer is not ganged sufficiently accurately, that is, the gains of the two sides are not maintained sufficiently similar as the control is turned; we have seen that under these circumstances the rejection ratio is adversely affected. Nevertheless the potentiometric method may be adapted successfully to achieve control in discrete steps by using a pair of potential dividers composed of precision resistors, and the range of control may be made as wide as desired in this manner (*Figure 12.28*). It is then usually necessary to 'fill in' between the steps with a second continuously variable control, of limited range only, which may be of the feedback type. To achieve gain control by

GAIN AND BALANCE CONTROLS

feedback R_K is split into two parallel parts and the cathodes are joined by a variable resistance (Figure 12.29). This usually involves the use of a negative HT supply, as in d.c. amplifiers. When R_T is reduced to zero the stage has a

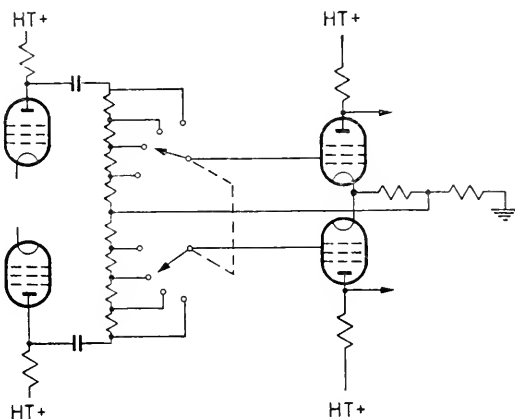


Figure 12.28

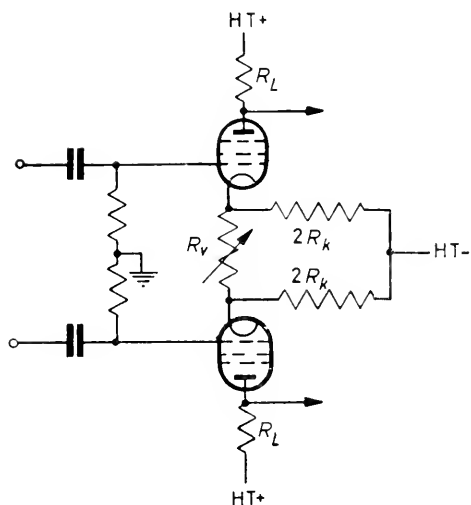


Figure 12.29

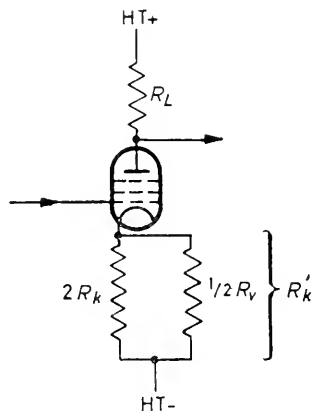


Figure 12.30

long tail, R_K , as before; and the anti-phase gain—if both outputs be used—is $g_m R_L$. As R_T is increased, cathode follower-like action becomes possible and the anti-phase gain falls. It is clear from the symmetry of the system that the mid-point of R_V must remain at a fixed potential and therefore as far as potential changes are concerned the equivalent circuit of each half is as that shown in Figure 12.30.

We have

$$\text{total gain} = \frac{A}{1 + AB}$$

DOUBLE-SIDED AMPLIFIERS

where

$$A = g_m(R_L + R_{K'})$$

$$B = \frac{R_{K'}}{R_{K'} + R_L}$$

and $R_{K'}$ is the effective resistance in the cathode circuit of each half

$$R_{K'} = \frac{\frac{1}{2}R_V \cdot 2R_K}{\frac{1}{2}R_V + 2R_K}$$

Of this the useful fraction is

$$\frac{R_L}{R_L + R_{K'}}$$

so the usable gain is

$$\frac{R_L}{R_L + R_{K'}} \cdot \frac{g_m(R_L + R_{K'})}{1 + g_m(R_L + R_{K'}) \left(\frac{R'_{K'}}{R'_{K'} + R_L} \right)} = \frac{g_m \cdot R_L}{1 + g_m R_{K'}}$$

This is plotted in *Graph 29* for a typical stage having $R_L = R_K = 100 \text{ k}\Omega$, and $g_m = 1 \text{ mA/V}$. It is clear that the control obtained is very non-linear if a wide range of gain control is required. However, if a threefold reduction is sufficient, a linear law $5 \text{ k}\Omega$ potentiometer may be calibrated in dB's without producing an immoderately non-linear scale.

Notice that the presence of R_V does not affect the operation of the stage to in-phase signals, for then both cathodes move together, no current flows in R_V and therefore its value is irrelevant. Notice also that the reduction in gain secured by this method must not be too great, otherwise, in-so-far as the anti-phase gain/in-phase gain ratio is being reduced, the overall amplifier rejection is adversely affected.

D.c. amplifiers

Potentiometric control of the gain of direct coupled amplifiers is complicated by the requirement that alteration of the gain control when the input to the amplifier is zero ought to produce no effect at the output, since zero times any number is still zero. Thus some kind of control based on *Figure 12.31* is no use because there is a potential gradient along the gain potentiometer, and manipulation of the gain control alters the bias on the second valve.

It is possible to devise several networks in which no standing current flows along the gain potentiometer; probably the simplest is *Figure 12.32*. The control is necessarily in discrete steps in order to preserve the accuracy of gain balance. R_1 and R_2 are compensated coupling potential dividers, and when the amplifier is in the resting state points A and B are at the same potential. Thus altering the stepped gain setting has no effect on the potentials supplied to the following grids. When an anti-phase signal is applied to the amplifier, current flows between A and B and a greater or lesser potential difference is tapped off according to the setting of the switches.

Remember that the effective anode load for each of the first stage valves in *Figure 12.32* is now $(R_L R_Z)/(R_L + R_Z)$ and therefore that the maximum gain will be reduced if this gain circuit is added to an existing design. If starting a design from scratch the effect is offset by choosing R_L about double

GAIN AND BALANCE CONTROLS

the usual value and making R_L and R_Z approximately equal. A higher value of HT positive is generally required.

Feedback gain control, to fill in between steps, is applied to d.c. amplifiers in exactly the same way as to a.c. coupled.

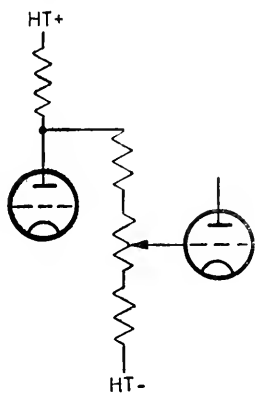


Figure 12.31

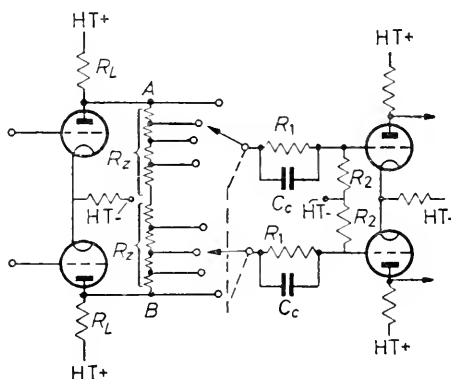


Figure 12.32

Gain balance

We have seen earlier that it is particularly important to arrange that the gain of the two halves of the *first* amplifying stage be made equal, and this is achieved by a differential gain control, i.e. one which increases the gain of one half and reduces it in the other.

This may conveniently be achieved by differential variation of the anode loads (Figure 12.33). If the valves can be selected for g_m 's within, say, 5 per

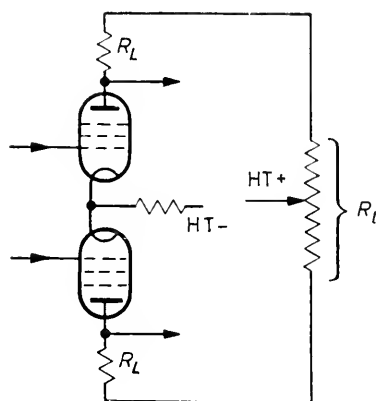


Figure 12.33

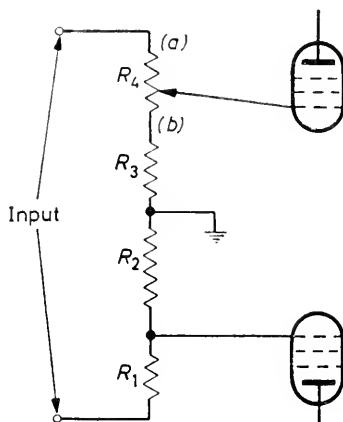


Figure 12.34

cent of a norm, then in the worst possible case the amplifier will have to compensate for a 10 per cent unbalance in g_m by providing up to 10 per cent unbalance in anode load. This it can do if R_1 is chosen 10 per cent of R_L , or a bit more for safety.

DOUBLE-SIDED AMPLIFIERS

Depending on the nature of the input circuits, it may be possible as an alternative to carry out gain balancing at the grids (*Figure 12.34*). R_1 and R_2 form a fixed potentiometer which reduces the input to the lower valve by a factor 0.9. If R_1 is a quarter of R_3 , then when the slider is in the middle the input to the upper valve is also reduced to 0.9. When the slider is at a the feed to the upper valve is 10 per cent high, and when at b it is 10 per cent low.

Static balance

This is a problem which occurs only in direct-coupled amplifiers, and refers to the requirement that when the difference of potential between the input terminals is zero the difference of potential between the output terminals is also zero. An a.c. coupled amplifier is self-adjusting in this respect, since in the absence of any input the potentials of the output terminals automatically return to that of the point P , usually earth, and so become equal to each other (*Figure 12.35*).

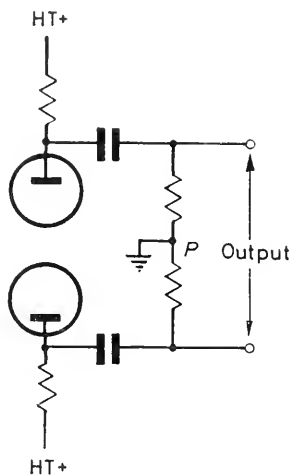


Figure 12.35

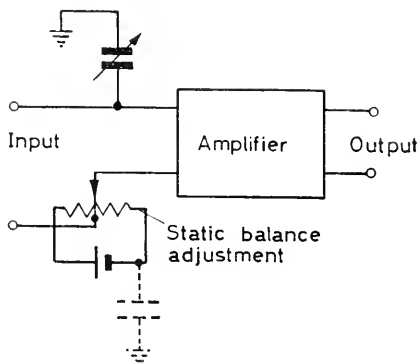


Figure 12.36

Lack of static balance occurs as a result of asymmetries in the circuits of three kinds:

- (1) lack of balance in the d.c. resistance of valve pairs in each stage when the grid potentials are equal;
- (2) lack of balance in the circuit resistors deliberately introduced to achieve gain balance;
- (3) numerous other effects, many of which fluctuate, to produce 'drift'.

It is possible but unwise to use direct-coupled amplifiers in a state of static unbalance. Quite a small amount of unbalance originating in an early stage is amplified by the later stages and the final stage may be operating under conditions of serious distortion—one valve passing grid current and the other cut off. Since the final stage is affected first, static-unbalance is best detected by a sensitive and high resistance voltmeter connected between the final stage anodes. If the final stage is operating in a statically balanced condition there is seldom cause to worry about the earlier ones.

GAIN AND BALANCE CONTROLS

As to how to correct for static unbalance, it is surprisingly difficult to devise methods for doing this which do not also upset the gain balance. One often sees arrangements such as the differential anode load being advanced as static balancing devices, and there is no question that they can be used as such; but if one then seeks some other method for doing one's gain balancing, it is sure to turn out that it throws the static balance off again. The only reliable method known to the author is to have a special source of variable e.m.f. with which to inject into the amplifiers at some appropriate point; a signal to correct the off-balance, for example *Figure 12.36*. This may be conveniently a small dry cell and potentiometer. These components have an appreciable capacitance to earth which affects the high frequency rejection properties of the amplifier unless corrected for by a similar capacitance between the other input terminal and earth. In order to prevent undue loss of high frequency response as a result of these capacitances, the input terminals of the amplifier should be fed from a pair of cathode followers. Since electro-physiological amplifiers are usually preceded by cathode followers anyway, this is not a disadvantage.

There remains one difficulty. If a single correcting voltage at the input is used to establish static balance at full gain, then when the gain setting is reduced the amplifier will in general go off balance. For example, consider *Figure 12.37*:

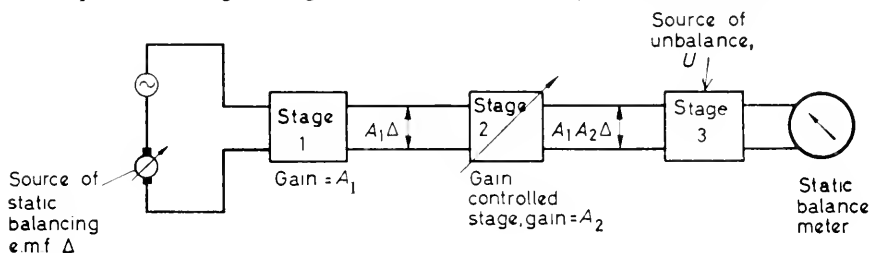


Figure 12.37

there is a lack of static balance, represented here by an equivalent unbalancing signal U in the stages following the gain control stage. At full gain this is offset by a static balancing e.m.f. Δ at the input, which emerges from the first stage as $A_1\Delta$ and from the second stage as $A_1A_2\Delta$. If the amplifier is statically balanced, the output of stage 3 is zero, so $A_1A_2\Delta$ must equal U , and Δ is adjusted until this is true. If now A_2 is altered, a new value of Δ will be required. This is unsatisfactory.

Fortunately there is another adjustment scheme which has the property that static balance correction may be carried out without upsetting the differential gain balance, and which only works at low gain settings. This is the differential cathode resistor. Consider *Figure 12.38*. At full gain $R_v = 0$ and the effect of manipulating the differential cathode potentiometer is to alter in opposite directions two resistances in parallel. The resistance of the parallel connection is at a very flat maximum when the potentiometer is at mid-setting, and the gains of the two halves are $g_{m1}R_1$ and $g_{m2}R_2$ independent of this setting. That is, operation of the control does not affect the gain balance. At low gain R_v is large and the effect produced by altering the cathode potentiometer is quite different, for it alters the bias on the valves

DOUBLE-SIDED AMPLIFIERS

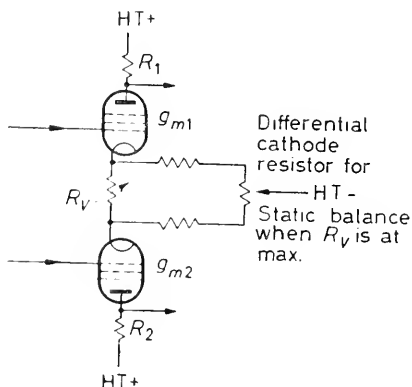


Figure 12.38

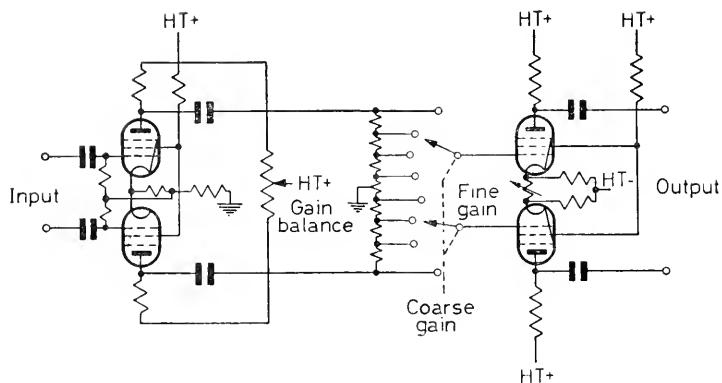


Figure 12.39

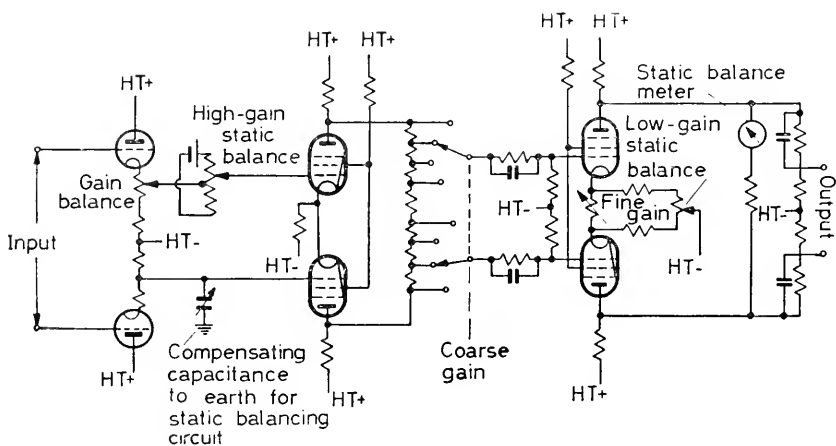


Figure 12.40

DIFFERENTIAL POWER AMPLIFIER

differentially and hence the anode currents, producing differential alteration of the anode potentials and therefore of the amplifier static balance condition. By combining the schemes of *Figures 12.36* and *12.38*, satisfactory static balancing becomes possible. If the amplifier is balanced when $R_f = 0$ (high gain) by the floating supply Δ , and when R_f is at maximum (low gain) by the differential cathode resistor, then it will remain balanced at intermediate settings of R_v . A possible outline design for an a.c. coupled differential voltage amplifier is given in *Figure 12.39*, and a d.c. version with cathode follower input in *Figure 12.40*.

DIFFERENTIAL POWER AMPLIFIER

This is an application to a power output stage of the property of the differential voltage amplifier of providing a push-pull output from a single-sided input. An a.c. coupled version is shown in *Figure 12.41*, and a d.c.

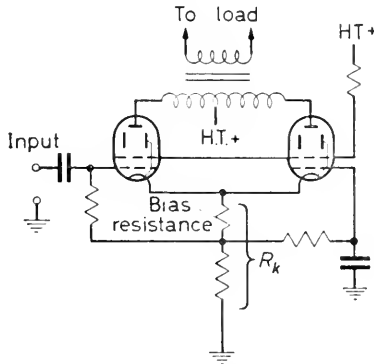


Figure 12.41

version in *Figure 12.42*. The advantage of the method is its simplicity in that no phase-splitter is required. The disadvantage is low power efficiency; the valves have to be operated in class A for the scheme to work, and for the

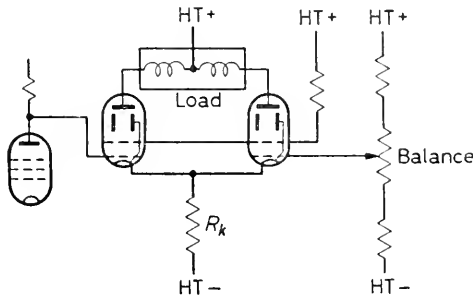


Figure 12.42

signal to be nearly equally divided between the valves, R_K should be large. Thus a standing current flows both in the valves and in R_K , whether there is any signal passing through the stage or not. An a.c. coupled cathode-coupled

DOUBLE-SIDED AMPLIFIERS

output stage is used in a commercial radiogram¹. A direct-coupled version is used in a Velodyne circuit².

PUSH-PULL POWER STAGES FOR LOADS WITHOUT A CENTRE-TAP

Here we are concerned mainly with direct-coupled loads, such as meters and some kinds of penwriters. Two possible schemes are shown in *Figure 12.43*.

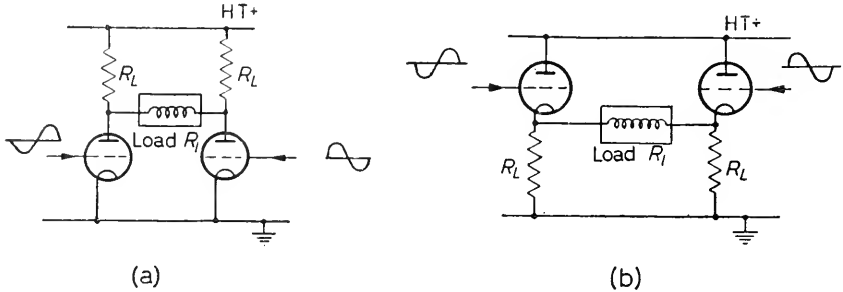


Figure 12.43

Figure 12.43a requires less input to the grids to secure a given load current, but *b* may be preferable in cases where the load has to be fed from a source of low impedance (as is sometimes necessary to get the correct electromagnetic damping with penwriters): neither of these circuits has very good power efficiency. Suppose as a result of a very large input signal one valve is cut off and the other becomes for all intents and purposes a short circuit, then the maximum possible load current which can flow in either case is $V/(R_L + R_L)$ (*Figure 12.44a* and *b*), so that we would like to have R_L small compared with

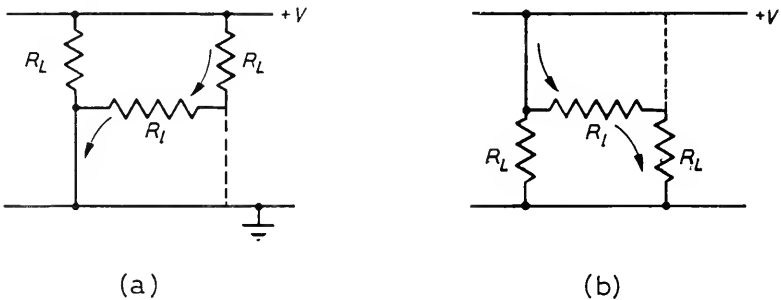


Figure 12.44

R_L . Also, the maximum voltage swing at the output is clearly obtainable when the no-signal potential of each load terminal is $+(V/2)$, for then the stage can theoretically operate until one terminal rises to $+V$ and the other falls to earth potential. It follows that the optimum resting current through each valve in either configuration is given by $\frac{1}{2}V/R_L$ and is small when R_L is large. Thus the requirement for a low standing HT current conflicts with the requirement for good maximum load current.

PUSH-PULL POWER STAGES FOR LOADS

A reasonable procedure for feeding penwriters is to make $R_L = R_L$. Then if $V = 250$ V and $R_L = 2,500 \Omega$, the standing current consumed by each valve is $125/2,500 = 50$ mA, and the maximum possible load current is 25 mA. To have to provide a continual 100 mA to a stage capable of supplying only 25 to the load is clearly rather wasteful. If the load is a meter, the maximum load current required is likely to be much less, and R_L can safely be made much greater than R_L .

Valve-bridge output stage.

The power inefficiency of the preceding system is overcome by the bridge scheme shown in Figure 12.45. These valves are arranged so that an increase in the current through V_1 and V_4 coincides with a decrease in the current through V_2 and V_3 . In the limiting case where V_1 and V_4 approximate to short circuits and V_2 and V_3 are nearly cut off the arrangement reduces to Figure 12.46 so that all the current consumed passes through the

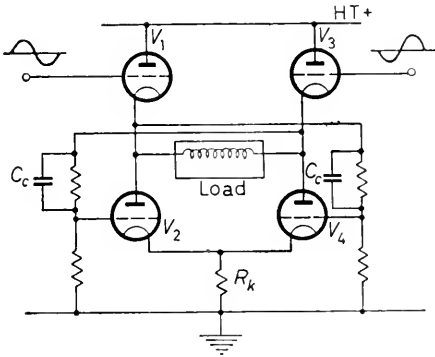


Figure 12.45

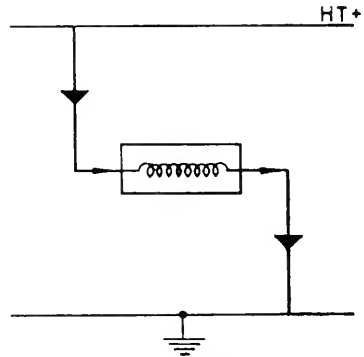


Figure 12.46

load. V_1 and V_2 act as cathode followers having V_2 and V_4 as cathode loads. V_2 and V_4 act as amplifiers having V_1 and V_3 as anode loads, and are biased by R_L such that HT is evenly divided between upper and lower valves. In order that upper and lower valves have equal effect on the currents flowing, the lower valve grids are fed from potential dividers. The division ratio is correct when a signal applied *in-phase* to the stage input terminals increases or decreases all the valve currents equally, such that no change of potential occurs either across the load, or between load and earth.

REFERENCES

- ¹ *Wireless World* 62 (1951) 389
- ² UTTLEY A. M. and WILLIAMS F. C. *J. Instn. elect. Engrs.* 93, Pt IIIA, No. 7 (1946) 1256

TUNED AMPLIFIERS

It often happens that amplifiers are required to amplify preferentially a certain frequency (acceptor amplifiers) or else to amplify all *except* a certain frequency (rejector amplifiers). In radio practice such amplifiers are the rule rather than the exception, and tuning is obtained by resonant circuits comprising inductance and capacitance. Q values of 100 or more are easily

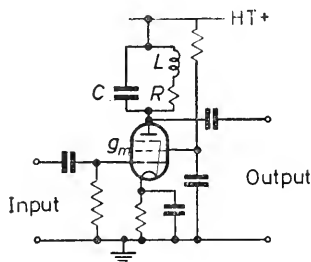


Figure 13.1

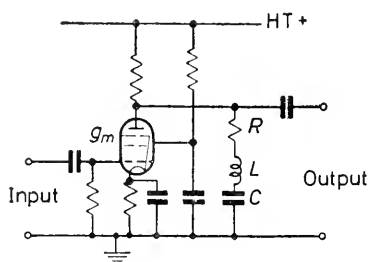


Figure 13.2

obtainable. Thus a simple acceptor amplifier comprises a pentode having for its anode load a parallel tuned circuit (*Figure 13.1*). The gain is $g_m |Z_p|$, and we saw in Chapter 5 that

$$\frac{|Z_p|}{\omega_o L} = \left\{ \frac{\left(\frac{\omega_o}{\omega}\right)^2 + Q^2}{1 + Q^2 \left(\frac{\omega_o}{\omega} - \frac{\omega}{\omega_o}\right)^2} \right\}^{1/2}$$

Since g_m is constant, the gain of the stage has the same form as $|Z_p|$ in *Graph 25*.

Similarly a rejector amplifier might have the form of *Figure 13.2*. The series tuned circuit is effectively in parallel with the load resistance and will have—at least near resonance—a much lower impedance than it. The gain will be substantially $g_m |Z_s|$ where

$$|Z_s| \cdot \omega_o C = \left\{ \frac{1}{Q^2} + \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)^2 \right\}^{1/2}$$

and the gain has the same form as $|Z_s|$ in *Graph 24*.

When we try to apply these circuits to the kind of frequencies commonly encountered in electrobiology we run at once into difficulties. Suppose it is required to build an attachment to an E.E.G. apparatus which selectively amplifies α rhythm at 10 c/s. Then ω is about 60 and the LC product

TUNED AMPLIFIERS

required = $1/\omega^2 = 1/3,600$ henry farads. The largest paper-insulated capacitors commonly available have a value of $8 \mu\text{F}$. If we employ one of these the inductance required

$$L = \frac{1}{3,600 \times 8 \times 10^{-6}} = 35 \text{ H}$$

$8 \mu\text{F}$ paper capacitors and 35 henry inductors are bulky and expensive components. As has been hinted before, inductors are a nuisance in electronic designs. They usually have to be home-made and they ought to be enclosed in expensive mumetal screening boxes.

An alternative approach lies through networks which use only resistances and capacitances. We have met the twin T , which clearly has a rejector characteristic, and a network exhibiting an acceptor characteristic is the band-pass filter for $a = 1$. These are clearly very unselective circuits.

The Q factor of inductance-capacitance tuned circuits may be variously defined. Up till now we have called it $\omega L/R$, or $(L/CR^2)^{1/2}$. If a response curve for a circuit of resonant frequency ω_0 is available, Q may be measured off as follows: find the points on the curve at which the response is $1/(2)^{1/2}$ of its maximum value. If the difference between the frequencies at which this occurs is $\delta\omega$, then $Q = \omega_0/\delta\omega$ (Figure 13.3). Strictly speaking this is only

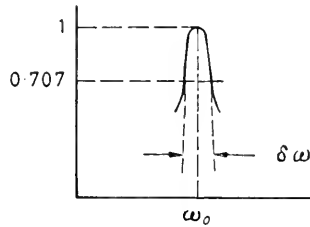


Figure 13.3

true when Q is quite large, but the relationship gives a very fair indication even for Q values less than 1. This gives us a basis for allotting an 'effective Q factor' to any resonance-like curve. Thus in *Graph 14* we see that when $a = 1$ the filter has $V_{\text{out}}/V_{\text{in}} = 0.37$ at 'resonance'. $1/(2)^{1/2}$ of this is 0.245, and we see the curve has these values of $V_{\text{out}}/V_{\text{in}}$ at $3/CR$ and $0.33/CR$.

Hence

$$Q = \frac{1/CR}{3/CR - 0.33/CR} = 0.375$$

Similarly, measurement from the twin T characteristic (*Graph 18*) reveals a Q factor of only about $\frac{1}{4}$. The output falls from its maximum value, 1, to $1/(2)^{1/2}$ at frequencies of approximately $4/CR$ and $0.25/CR$. These values do not look very good beside the Q 's of 100 or so obtainable from an inductance-capacitance tuned circuit.

There are three approaches by which we can improve matters. They involve

TUNED AMPLIFIERS

the incorporation of frequency-selective networks in valve circuits possessing feedback.

ACCEPTOR AMPLIFIER—POSITIVE FEEDBACK METHOD

We have seen that the application of negative feedback round an amplifier tends to reduce the effect of variations of amplifier gain with frequency (or with anything else). It is therefore intuitive that the employment of positive feedback will serve to exaggerate them. This is exactly what we want to improve the selectivity of our band-pass circuit.

Suppose we have an amplifier of gain A and insert into it somewhere in the signal path a band-pass filter whose transmission factor we will for the present designate $f(\omega)$. Then the gain without feedback becomes $Af(\omega)$ and if we now feed back positively a fraction of the output B (Figure 13.4) then

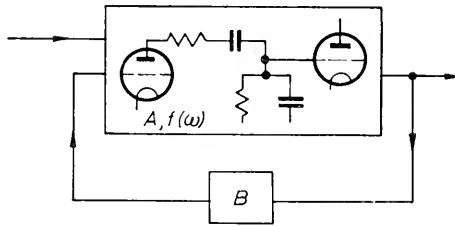


Figure 13.4

the gain with feedback is

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{Af(\omega)}{1 - Af(\omega)B}$$

which simplifies to

$$\frac{A}{\frac{1}{f(\omega)} - AB}$$

Now we can see from the results in Chapter 3 that for the R-C band-pass filter having $a = 1$,

$$f(\omega) = \frac{1}{3 + j\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)}$$

where $\omega_o = \frac{1}{CR}$

Hence $\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{A}{3 + j\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right) - AB}$

and $\left|\frac{V_{\text{out}}}{V_{\text{in}}}\right| = \frac{A}{\left\{(3 - AB)^2 + \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}\right)^2\right\}^{1/2}}$ *

This is clearly a curve which passes through a maximum at a 'resonant' frequency $\omega = \omega_o$ and at resonance the gain will be just $A/(3 - AB)$. It

remains to find an expression for the Q . When the gain has fallen to $1/(2)^{1/2}$ of its maximum value, it will be

$$\frac{A}{(2)^{1/2}(3 - AB)}$$

$$= \frac{A}{\{2(3 - AB)^2\}^{1/2}}$$

but this is the value of the starred equation when $\{(\omega/\omega_o) - (\omega_o/\omega)\} = \pm(3 - AB)$, from which we can deduce the upper and lower values of ω (let them be ω_1 and ω_2) at which the response is down to $1/(2)^{1/2}$ of its maximum.

For let $\omega_o + (\delta\omega/2) = \omega_2$, and $\omega_o - (\delta\omega/2) = \omega_1$,

so that
$$\frac{\omega_2}{\omega_o} = \frac{\omega_o + \frac{\delta\omega}{2}}{\omega_o}$$

and
$$\frac{\omega_o}{\omega_2} \approx \frac{\omega_1}{\omega_o} = \frac{\omega_o - \frac{\delta\omega}{2}}{\omega_o}$$

Then $\{(\omega_2/\omega_o) - (\omega_o/\omega_2)\}$, which is the positive case of $\{(\omega/\omega_o) - (\omega_o/\omega)\}$

$$= \frac{\delta\omega}{\omega_o}, \quad \text{but} \quad \frac{\delta\omega}{\omega_o} = \frac{1}{Q} \quad \therefore \quad Q = \frac{1}{3 - AB}$$

Thus, provided $A > 3$, a degree of amplification difficult to fail to achieve, any desired selectivity may be had by appropriate choice of B .

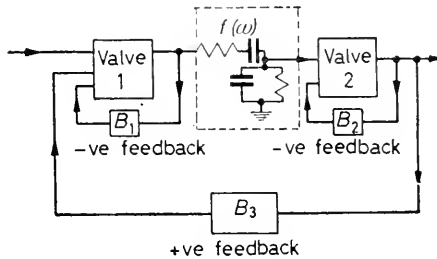


Figure 13.5

Although the minimum amplification required for indefinitely large Q is so low, it is usual to employ two valves in acceptor amplifiers of this kind for three reasons:

- (1) it facilitates the application of the feedback;
- (2) the spare gain can be lost in negative feedback loops *round the valves only*—this will stabilize A and thus the stability of the system as a whole (Figure 13.5);
- (3) where considerable gain is required but not a large Q , the amplification has to come from the valves and not from the positive feedback.

TUNED AMPLIFIERS

In *Graph 30* some performance curves for tuned amplifiers of the positive feedback type are plotted. A is arbitrarily fixed at 100. *Figure 13.6* shows a possible circuit for a positive feedback tuned amplifier.

Remember that the practical value of R' is the calculated value less the resistance looking back into V_1 , R_L in parallel with r_a . Similarly the practical

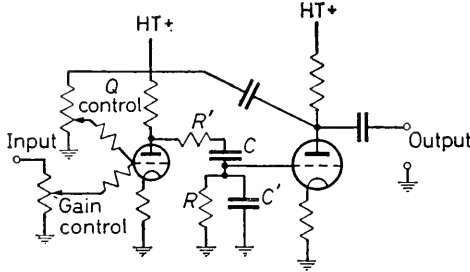


Figure 13.6

value of C' is the theoretical value less the stray and input capacitance of V_2 . Notice the un-bypassed cathode bias resistances which stabilize the gains of the two stages. Practical examples of circuits of this type have been published by Morris and Dawe¹.

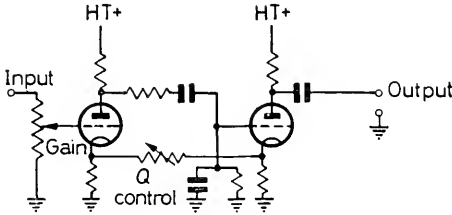


Figure 13.7

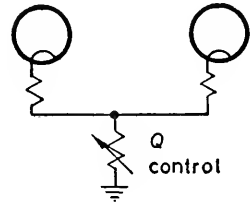


Figure 13.8

Alternatively, positive feedback may be had by coupling back from the second valve cathode to the first. This may be achieved by the delta resistance network of *Figure 13.7*, or the equivalent T form (*Figure 13.8*). A design employing this principle has been published by Beattie and Conn².

ACCEPTOR AMPLIFIER—NEGATIVE FEEDBACK METHOD

Whereas in the positive feedback method we include a band-pass filter in the amplifier, in the negative feedback method we include a null-transmission filter—most conveniently a parallel T —in the feedback loop (*Figure 13.9*). Away from the null-frequency of the parallel T there is 100 per cent negative feedback with a consequent overall gain of only 1. At the null-frequency, the feedback disappears and the gain is A .

More precisely we have

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{A}{1 + AB}$$

B is the transmission characteristic for a parallel T , which we have seen is

$$\frac{1}{1 - \frac{4j}{\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}}}, \text{ where } \omega_o = \frac{1}{CR}$$

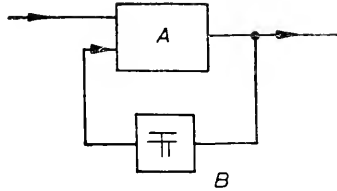


Figure 13.9

Thus

$$\begin{aligned} \frac{V_{\text{out}}}{V_{\text{in}}} &= \frac{A}{1 + A \left\{ \frac{1}{1 - \frac{4j}{\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}}} \right\}} \\ &= \frac{[A (\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}) - 4j]}{(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega})(1 + A) - 4j} \\ \left| \frac{V_{\text{out}}}{V_{\text{in}}} \right| &= A \left\{ \frac{\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 + 16}{(A + 1)^2 \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 + 16} \right\}^{1/2} \end{aligned}$$

This expression is, naturally enough, equal to A at $\omega = \omega_o$ and smaller elsewhere. To find the Q of the 'resonance' we consider, as before, the points at which the gain has fallen to $A/(2)^{1/2}$. At these points we have

$$\frac{A}{(2)^{1/2}} = A \left(\frac{1}{2} \right)^{1/2} = A \left\{ \frac{\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 + 16}{(A + 1)^2 \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 + 16} \right\}^{1/2}$$

$$\text{whence } 2 \left[\left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 + 16 \right] = (A + 1)^2 \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 + 16$$

$$\therefore ((A + 1)^2 - 2) \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 = 16$$

TUNED AMPLIFIERS

With any reasonable amplifier $(A + 1)^2$ will be much greater than 2

$$\therefore \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right)^2 \approx \frac{16}{(A + 1)^2}$$

$$\therefore \left(\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega} \right) \approx \pm \frac{4}{A + 1}$$

As we saw in the positive feedback acceptor amplifier, if we take the upper of the $1/(2)^{1/2}$ points ω_2 the $n\{(\omega_2/\omega_o) - (\omega_o/\omega_2)\}$ is positive and equal to $\delta\omega/\omega_o$.

$$\therefore \frac{\delta\omega}{\omega_o} = + \frac{4}{A + 1}$$

but

$$\frac{\delta\omega}{\omega_o} = \frac{1}{Q}$$

$$Q = \frac{A + 1}{4}$$

Some performance curves for acceptor amplifiers of this type are shown in *Graph 31*. It is clear that, amplifier gain for amplifier gain, the positive feedback system can give greater selectivity; also only two elements have to be adjustable to vary the frequency in the positive feedback case. On the other hand, where constancy of gain at the resonant frequency is important the negative feedback scheme is probably safer.

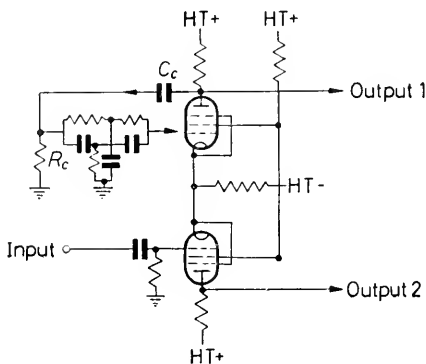


Figure 13.10

In designing circuits of this type it has to be borne in mind that the parallel T equation used here assumes the network to work into an infinite impedance. It ought therefore to feed back straight to a valve grid, itself unconnected to anything else, and the question then arises as to where to feed in the input. An elegant solution is to make use of a long-tailed pair, which has two input grids. The circuit has the appearance of *Figure 13.10*, and has the incidental advantage of providing balanced outputs if required. R_c can be comparable in value with the anode load resistors, and $C_c R_c$ should be chosen to produce

MIXED FEEDBACK METHOD—PHASE SHIFT ACCEPTOR AMPLIFIER

negligible phase shift at the resonant frequency, and the twin T resistances should be about ten times higher than the anode loads and R_c in order not to interfere with the stage gain without feedback.

A simpler and less efficient circuit is shown in *Figure 13.11*. The input

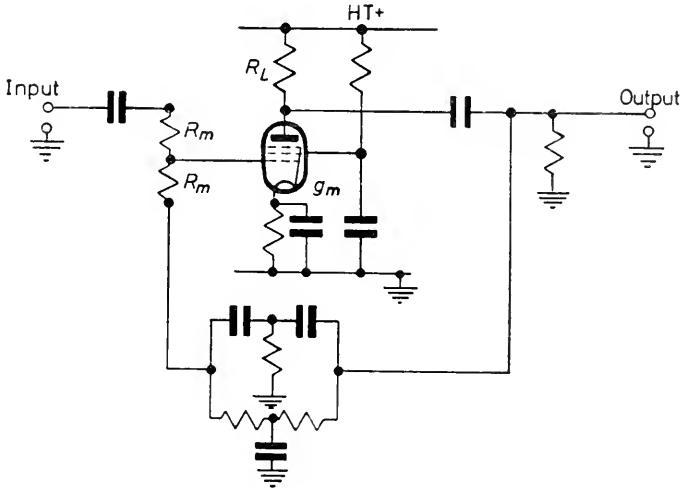


Figure 13.11

and feedback voltage are mixed by resistors R_m . These may conveniently be made equal and as large as possible consistent with the prevention of significant phase shift occurring between the output of the parallel T and the valve grid as a result of the valve input capacitance. Fleischer³ suggests that if the resistance seen looking into the mixing resistors from the parallel T is more than $2\frac{1}{2}$ times the resistance looking back into the parallel T , the loading effect on the parallel T will not be too serious.

If the mixing resistors are equal the inputs to the valve will be half the input signal and half the feedback signal. The system will therefore behave according to theory except that it is as if the valve gain had been halved. Thus with this circuit the value for A is $(g_m R_L)/2$ and not $g_m R_L$.

A circuit of the *Figure 13.10* type is published by Sowerby⁴.

MIXED FEEDBACK METHOD—PHASE SHIFT ACCEPTOR AMPLIFIER

This might be described as the poor man's tuned amplifier, as an indefinitely large Q may be obtained with only 1 valve. The feedback is basically negative, in-so-far as the feedback loop is from the anode back to the grid (*Figure 13.12*). There is in the forward path a 3-section R-C high-pass filter, each section having the same turn-over frequency, and a tapering factor between sections of a . The method of operation is as follows.

At very low frequencies the attenuation in the filter is so great that the output is very small. At very high frequencies the filter offers no attenuation and no phase shift, so that if a fraction of the output B is fed back to the

TUNED AMPLIFIERS

input, by appropriate operation of the potentiometer—the Q control—then the overall gain at high frequencies is $A/(1 + AB)$. At some intermediate frequency ω_r the phase shift per filter section becomes 60 degrees and the overall phase shift is therefore 180 degrees. The feedback is now positive.

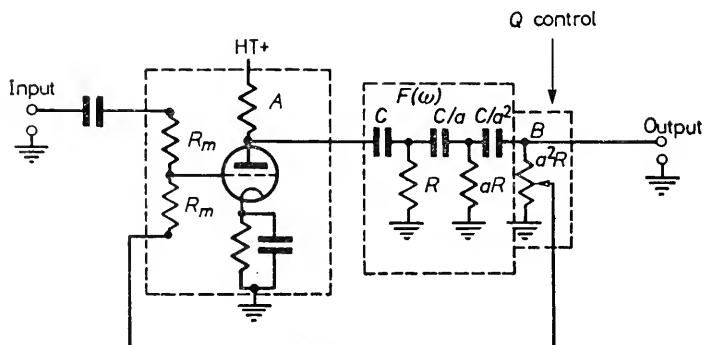


Figure 13.12

The overall gain rises to a maximum of $A/(1 - ABF(\omega_r))$ where $F(\omega)$ is the transmission factor of the filter.

The value of $F(\omega_r)$ depends on the filter tapering factor a . If all the sections are identical, $a = 1$, it can be shown that $|F(\omega_r)|$ is $1/29$ and $\omega_r = 1/(6)^{1/2}CR$. In our analysis we shall assume that the taper factor is sufficient for sections not to load their predecessors significantly, say $a = 10$. If this is true, it is easily seen that when the phase shift per section is 60 degrees, the output from each section is half the input and for three sections $F(\omega_r) = 1/8$. Thus for a phase shift tuned amplifier employing tapered sections the overall gain (and therefore the Q) rises to infinity when the denominator of the gain expression equals 0, i.e. $1/8 AB = 1$. Thus if B is made 1 the gain required of the amplifier in the dotted box of Figure 13.12 is only 8. If all the sections are identical the gain required is 29, but due to the mixing resistances, the gains required of the valve must be 16 and 58 respectively.

To show the performance more precisely we have for a single RC high-pass section

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{R}{R + \frac{1}{j\omega C}} = \frac{j\omega CR}{j\omega CR + 1}$$

So for 3 highly tapered sections

$$\frac{V_{\text{out}}}{V_{\text{in}}} = F(\omega) = \left(\frac{j\omega CR}{j\omega CR + 1} \right)^3 = \frac{-j \left(\frac{\omega}{\omega_o} \right)^3}{\left[1 - 3 \left(\frac{\omega}{\omega_o} \right)^2 \right] + j \left[3 \frac{\omega}{\omega_o} - \left(\frac{\omega}{\omega_o} \right)^3 \right]}$$

where
$$\omega_o = \frac{1}{CR}$$

Thus

$$\begin{aligned} \text{Overall gain} &= \frac{Af(\omega)}{1 + Af(\omega)B} = \frac{A}{\frac{1}{f(\omega)} + AB} \\ &= \frac{A}{\left[1 - 3\left(\frac{\omega}{\omega_o}\right)^2\right] + j\left[3\frac{\omega}{\omega_o} - \left(\frac{\omega}{\omega_o}\right)^3\right] - j\left(\frac{\omega}{\omega_o}\right)^3} + AB \end{aligned}$$

And

$$|\text{Overall gain}| = \frac{A}{\left\{\left[\left(\frac{\omega_o}{\omega}\right)^3 - 3\frac{\omega_o}{\omega}\right]^2 + \left[3\left(\frac{\omega_o}{\omega}\right)^2 - (1 + AB)\right]^2\right\}^{1/2}}$$

This has a maximum near $(\omega_o/\omega)^3 = 3(\omega_o/\omega)$, i.e. where $\omega = \omega_o/(3)^{1/2}$. By differentiating the above with respect to ω and equating to zero, it can be shown that the exact frequency at which the maximum occurs is given by:

$$\omega = \omega_r = \omega_o \cdot \frac{1}{(\sqrt{2AB} - 1)^{1/2}}$$

Evidently ω_r approaches $\omega_o/(3)^{1/2}$ as AB approaches 8. It can also be shown that, at ω_r ,

$$|\text{Overall gain}| = \frac{A}{2(2AB)^{1/2} - AB}$$

An expression for the Q , based on the points at which the ‘resonant’ gain is reduced by the factor $1/(2)^{1/2}$, has so far eluded the writer; some very cumbersome equations seem unavoidable, but we can get at least a measure of the selectivity of the amplifier by remarking the resemblance between the curves of *Graph 32* and those for the voltage across the inductance in a series L-C-R circuit (*Figure 5.41*). In the latter case Q was seen to be defined by

$$\frac{\text{Output at resonant frequency}}{\text{Output at frequencies much above resonance}}$$

Defining the selectivity of the amplifier in an analogous manner, and calling it \bar{Q} ,

$$\begin{aligned} \bar{Q} &= \frac{A}{\frac{2(2AB)^{1/2} - AB}{A}} \\ &= \frac{1 + AB}{2(2AB)^{1/2} - AB} \end{aligned}$$

Some performance curves for this type of amplifier are shown in *Graph 32*.

The design of these amplifiers has been more fully discussed by Shaw⁵,

TUNED AMPLIFIERS

who points out the difficulty in trying to use more than rather moderate tapering factors. The impedance of the first filter section ought to be appreciably larger than the anode load; the impedance of the last filter section ought to be appreciably less than the mixing resistances: hence the limitation on a .

It is not necessary to use a high-pass filter but it is quite useful to do so for filters peaking at frequencies below a kilocycle or so, because the first capacitor forms the HT blocking capacitance. At high frequencies the stray capacitance of the output circuit becomes significant and it is better to use low-pass sections, the last section being designed bearing the presence of the stray capacitances in mind. The effect is to reverse the gain/frequency characteristic left-right.

REJECTOR AMPLIFIER

Where the rejector characteristic of a simple parallel T filter is insufficiently sharp, improved performance may be had by following it with an acceptor amplifier of appropriate characteristic (*Figure 13.13*). The appropriate

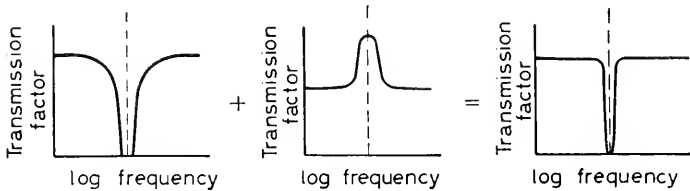


Figure 13.13

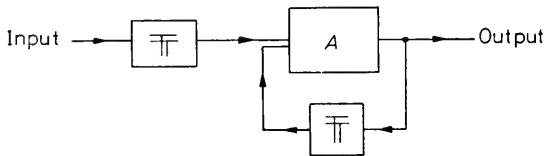


Figure 13.14

characteristic here is given by an acceptor of the negative feedback type, i.e. one having another parallel T in the feedback loop (*Figure 13.14*).

The characteristic of the parallel T is given by $1/(1 - 4j\sigma)$

where

$$\sigma = \frac{1}{\frac{\omega}{\omega_o} - \frac{\omega_o}{\omega}} \quad \text{and} \quad \omega_o = \frac{1}{CR}$$

as we have seen. Further, the characteristic of the acceptor amplifier is given, as we have also seen, by

$$\frac{A}{1 + A \frac{1}{1 - 4j\sigma}}$$

REJECTOR AMPLIFIER

The transmission characteristic for the pair in tandem is thus given by

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{A}{1 + A \frac{1}{1 - 4j\sigma}} \cdot \frac{1}{1 - 4j\sigma} = \frac{A}{1 - 4j\sigma + A}$$

If $A \gg 1$

$$\frac{V_{\text{out}}}{V_{\text{in}}} \approx \frac{1}{1 - 4j\frac{\sigma}{A}}$$

This, however, is of the same form as the expression for the parallel T alone, except for an A -fold expansion of the scale of frequency. The effect is thus to squeeze the parallel T characteristic to $1/A$ of its usual width.

REFERENCES

- ¹ MORRIS, G. W. and DAWE, P. G. M. A bandpass filter for low frequencies *Electron. Engng.* 25 (1953) 365
- ² BEATTIE, J. R. and CONN, G. K. T. A simple low-frequency amplifier *Electron. Engng.* 25 (1953) 299
- ³ VALLEY, G. and WALLMAN, H. *Vacuum Tube Amplifiers* New York; McGraw-Hill
- ⁴ SOWERBY, J. Mc G. Electronic circuitry *Wireless World* 56 (1950) 223
- ⁵ SHAW, J. C. The design of phase shift oscillators and filters *Electron. Engng.* 28 (1956) 216

SINE WAVE OSCILLATORS

Oscillators producing a sine wave output occur frequently in electrobiological work, important applications being the production of time markers from an alternating source of accurately known frequency, and the generation of test signals of variable known amplitude and frequency for measuring the performance of amplifiers and recording gear. There are three important types: (1) LC oscillators; (2) Crystal oscillators; and (3) RC oscillators.

LC OSCILLATORS

LC oscillators are suited to the production of frequencies above, say, 1 kilocycle (where the amount of capacitance and inductance required to achieve resonance is not inconveniently great) provided that the range of frequency adjustment required does not exceed about 2 to 1. The reason for this limitation is that the amplitude of oscillation depends on the Q of the resonant circuit $= (L/CR^2)^{1/2}$. Thus for constant amplitude over the range of frequencies required the ratio L/C ought to be held constant, while the product LC (remember resonant frequency equals $1/LC$ for series tuned circuit and approximately $1/LC$ for parallel tuned circuit) is varied. This requires ganging of variable inductors and capacitors, rather an elaborate business.

If a parallel tuned circuit be used as the load for a valve whose anode current contains an alternating component of frequency ω and of instantaneous value i , and if the tuned circuit is resonant at the frequency ω , the valve sees an a.c. load which is pure resistive and is equal to L/CR , and a circulating current jQi flows round the tuned circuit. If a second inductance be loosely mutually coupled to the tuned circuit inductance, then the e.m.f. induced in the second inductance is $j\omega M(jQi) = -\omega MQi$. If this e.m.f. is applied to the grid of the valve—due regard being paid to the sense of the connections—the feedback is positive and without phase shift, which means that if M , Q and the valve gain are great enough the circuit will generate continuous oscillations at frequency ω (*Figure 14.1*).

Now this is a feedback system of the kind we have become familiar with, where the loop gain is clearly $AB = g_m \omega MQ$ for pentodes and $\mu / \{(L/CR) + r_a\} \omega MQ$ for triodes. The criterion for oscillation is that $AB = 1$. If AB becomes < 1 the oscillations die away and if $AB > 1$ they will—theoretically—grow to infinity. Thus the production of oscillations of constant amplitude requires that AB should be held exactly equal to 1 at all times. It is in the nature of practical things that this sort of state of affairs, even if it could be temporarily established, would not persist for very long and in general sine wave oscillators require some sort of automatic regulating device to maintain the amplitude of the output.

In the LC oscillator this is simply achieved by the addition of a capacitance and a high resistance or 'grid leak', and leaving out any form of grid bias

LC OSCILLATORS

supply (Figure 14.2). On first switching on, C_2 is uncharged, the valve receives no mean bias and the loop gain is arranged to be amply large for oscillations to begin and to grow. When the point A swings positive the grid will also be carried positive and grid current will flow into C_2 and charge it.

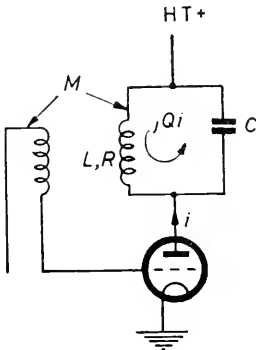


Figure 14.1

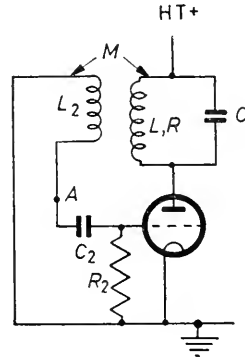


Figure 14.2

The effect of this charging will be such as to cause the grid to become less positive, as shown in Figure 14.3. The time constant $C_2 R_2$ is chosen so that it would take much longer than the time of one cycle of oscillation for the charge on C_2 to leak away via R_2 *. Thus when A swings negative again C_2 remains



Figure 14.3

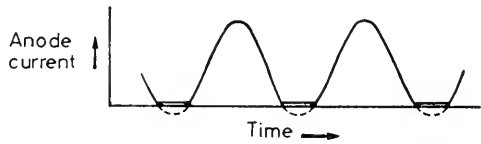


Figure 14.4

charged to approximately the peak voltage developed across L_2 . Thus the voltage to which C_2 charges represents a steady bias on the valve of magnitude proportional to the amplitude of oscillation.

As the oscillations grow a point is reached at which, on the extreme negative excursions of grid voltage, the valve is cut off altogether (Figure 14.4). It is now clear that the form of the anode current is of the form of a fundamental wave of frequency ω plus harmonics. More important, as a result of the clipping of the waveform, a given proportional increase in the valve input can no longer produce a commensurate increase in anode current swing, that is, the effective valve gain is reduced. The amplitude of oscillations is therefore automatically stabilized, for any tendency to increase further is offset by a further reduction in effective valve gain, and vice versa.

The equilibrium level of oscillation and the form of the anode current wave depend on the magnitude of Q , M and the valve gain. If this product is barely sufficient to maintain oscillation the output will be small and the anode current might have the form of Figure 14.5a. This is clearly substantially class

* But see under Blocking Oscillators

SINE WAVE OSCILLATORS

A operation, which the reader will remember is associated with poor efficiency because of the wasteful steady component of anode current. If now the Q - M gain product is increased—most conveniently by increasing M —the oscillations will be larger and the anode current might resemble *Figure 14.5b*.

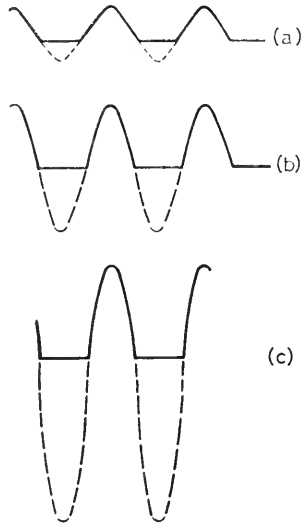


Figure 14.5

This is class B working, in which the steady component of anode current has been eliminated, leading to much better circuit efficiency. Finally, if M is made any larger, the anode current has the form of *Figure 14.5c*, in which the current flows only in short bursts. This is class C working, in which the valve gives itself a mean bias *beyond* cut off. This is the most efficient mode of working of all.

At this point it may be objected that the output of an oscillator having an anode current of the form of *Figure 14.5c* can surely not be very sinusoidal. The answer is, it all depends on the Q of the tuned circuit. If the Q is 100, then the fundamental component of the anode current will be magnified 100 times whereas the harmonics will not, with the result that the voltage wave across the tuned circuit is nevertheless of substantially pure sine form. If the Q is poor, the purity of the output will also be poor.

Hartley oscillator

The circuit so far discussed is not in fact much used; in practice two coils are not necessary, as the same effect may be had by employing a tuned circuit with a tapping point. If the tapping point is on the inductance the circuit is called the Hartley oscillator. The version in *Figure 14.6* is called the 'series fed' Hartley oscillator because part of the tuned circuit is in series with the HT supply to the valve. The 'shunt fed' Hartley circuit appears in *Figure 14.7*. It has the advantage, in some applications, that the tuned circuit is at earth potential so far as direct voltages are concerned. C_3 is required to block the HT and exhibit negligible reactance at the oscillation frequency, while the

LC OSCILLATORS

'choke' inductance is required to exhibit a high reactance at the oscillation frequency so that the a.c. load seen by the valve is still substantially the tuned circuit.

In both series and shunt-fed Hartley circuits the position of the tap determines the amount of feedback to the grid and hence the level of oscillation and

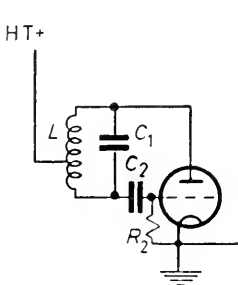


Figure 14.6

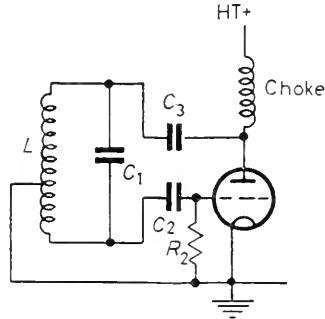


Figure 14.7

the class of valve operation. Starting at the bottom, moving the tap up the coil increases the feedback, and the loop gain up to a maximum; thereafter the loop gain falls again because the grid leak, in parallel with the valve input resistance, is reflected across the tuned circuit by autotransformer action as a progressively lower resistance, thus damping it and reducing the Q . The best position for the tap is usually found by experiment; about 1/5 of the way up is reasonable for a first guess.

Colpitts oscillator

In the Colpitts circuit (Figure 14.8) which must necessarily be shunt-fed, the tuned circuit capacitance is tapped instead of the inductance. If the total tuning capacitance required is C , then the capacitances C_a and C_b must be related to C by $C = (C_a C_b) / (C_a + C_b)$, and if the equivalent distance up the inductance for the tapping point is to be $1/n$ th, then $n = C_b + C_a / C_a$.

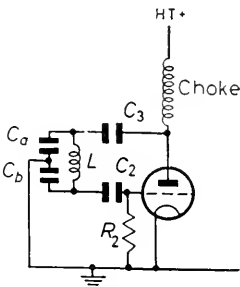


Figure 14.8

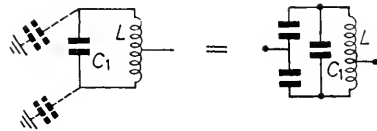


Figure 14.9

The Colpitts circuit requires one extra component and may be difficult to arrange if the frequency is to be adjusted by varying the total tuning capacitance. Its exponents claim for it an advantage over the Hartley oscillator which runs roughly as follows: In a Hartley oscillator there are in effect two

SINE WAVE OSCILLATORS

tapping points: one, the intentional one on the inductance, and another unintentional one of the Colpitts type in virtue of the two stray capacitances between the ends of the tuned circuit and earth (*Figure 14.9*). Since the two taps are unlikely to be at the same effective distance up the tuned circuit, the oscillator is liable to peculiar behaviour due to, as it were, not knowing which tap it is supposed to be using. The author has never experienced this difficulty and has so far always used the Hartley circuit with satisfactory results. Probably the Colpitts circuit becomes necessary only at frequencies much higher than those usually of interest to electrobiologists.

The Hartley and Colpitts oscillators are the important LC oscillators and are the only ones with which it is proposed to deal. We have now to attend to how to lead out the oscillations the circuit is generating.

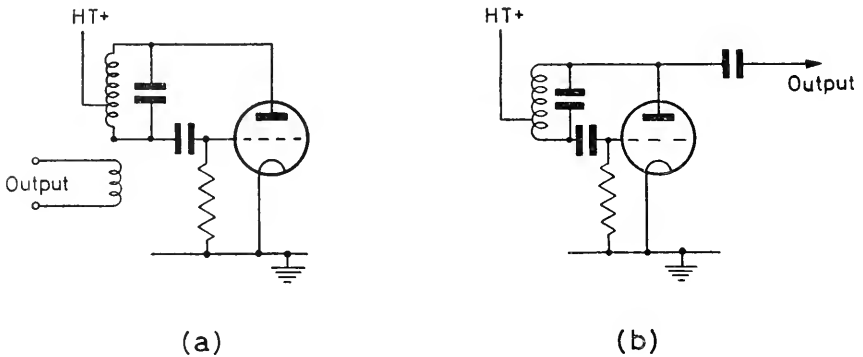


Figure 14.10

If the requirements of the oscillator in terms of frequency and amplitude stability are not too stringent, energy may be removed from the tuned circuit by an extra inductively coupled winding, or possibly even a lead taken—via a blocking capacitance—direct from the anode (*Figure 14.10*). A typical example is the erasing oscillator of magnetic tape recorders, where the

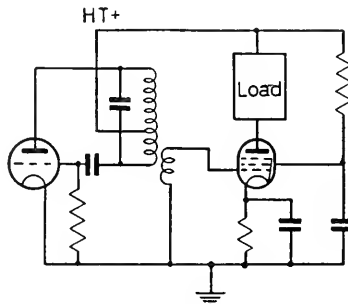


Figure 14.11

amplitude of the erase signal can be any size provided it is sufficient to eradicate previous recordings, and where any frequency will do provided it is several times the highest frequency that can be recorded. If the output is taken from the anode it is not difficult to get a peak-to-peak swing of nearly

LC OSCILLATORS

twice the HT voltage, which is large enough for most purposes. The output impedance, or impedance 'looking back' into the oscillator, will be the dynamic resistance of the tuned circuit in parallel with the r_a of the valve. Other output voltages and impedances may be had by inductive coupling and correct choice of the number of turns on the coupling winding, according to transformer theory.

The snag with simple schemes such as these is the upsetting effect upon the oscillator performance caused by variations in the nature of the load. If the latter is resistive it will affect the Q of the tuned circuit and the amplitude of the oscillation*. If the load is also reactive it will affect the frequency of oscillation as well. What is wanted is some form of 'buffer' or isolating device between the two: this may take the form of another valve, preferably a pentode (*Figure 14.11*), or resort may be had to a useful scheme known as 'electron coupling'.

Electron coupled oscillator

The electron coupled circuit is shown in its most simple form in *Figure 14.12*. It will be seen that a pentode is employed whose control grid and screen are

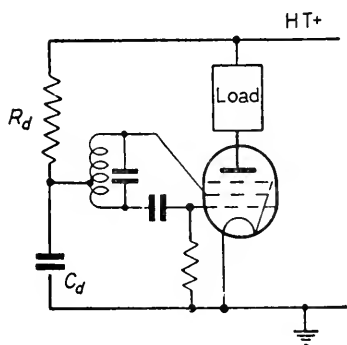


Figure 14.12

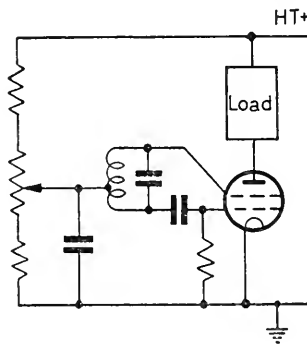


Figure 14.13

used to form a triode Hartley oscillator. R_d and C_d supply a direct decoupled screen potential. The suppressor is earthed in the usual manner, and the load is connected between anode and HT. Thus the oscillating triode section is screened from the anode and the effects of diverse anode loads by the suppressor. A small proportion of the electron stream passes to the screen and serves to sustain oscillation, the rest passes through to the anode to develop power in the load. The anode current is far from sinusoidal and a further tuned circuit usually constitutes the load in order to recover a pure waveform.

The pentode electron coupled oscillator frequency is thus independent of the load, but unfortunately it is not independent of variations in the HT supply. In the simplified description of LC oscillator operation, the tuned circuit was merely said to be 'resonant' but nothing was said about precisely how the resonant frequency is found. It is certainly not far from $\omega = (1/LC)^{1/2}$ and for practical purposes this is sufficient, since LC will be made adjustable and will be set up to give the frequency required. However, it is important to

* Also the frequency, slightly.

SINE WAVE OSCILLATORS

bear in mind that the *exact* expression for the resonant frequency for a tuned circuit performing sustained oscillations under the action of a valve contains terms involving the valve parameter r_a . These expressions are different for the various oscillator circuits or the two-coil circuit discussed

$$\omega_{\text{res}} = \left\{ \frac{1}{LC} \left(1 + \frac{R}{r_a} \right) \right\}^{1/2}$$

and we need not concern ourselves with them except to bear in mind that for constant ω , r_a should be constant. With tetrodes it is a fortunate fact that the effects of raising the anode and screen supply voltages produce opposite effects on the r_a , the screen effect being preponderant. Thus if the screen be fed from a suitable fraction of the HT supply it is possible to secure—at least over a range of HT variations—a constant r_a . Thus a tetrode circuit to use this phenomenon might have the form of *Figure 14.13*. The oscillator circuit is supplied from a potentiometer across the HT supply and the potentiometer is adjusted until the required frequency stability is obtained. Unfortunately we lose with this circuit the valuable property of isolation between oscillating circuit and load, since in the absence of the suppressor there is capacitance coupling between anode and screen. This difficulty is overcome in the following manner.

Consider once more the simple Hartley circuit of *Figure 14.6*. There is no particular reason why earth should be connected to the valve cathode. Let us connect it instead to the anode, and re-draw the circuit as in *Figure 14.14*. The anode circuit of the valve is now *ABCDEF*. It does not matter whereabouts in this circuit the HT supply appears, so let us remove it from *EF* and put it between *BC*. We now have *Figure 14.15*. This is still a Hartley oscillator, though possibly in unfamiliar form.

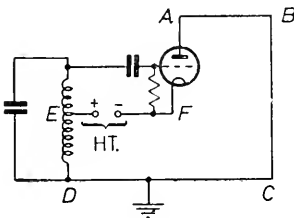


Figure 14.14

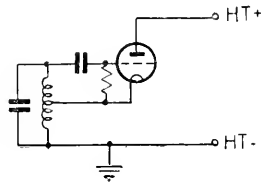


Figure 14.15

If now this circuit be used as the triode oscillator section of an electron coupled tetrode circuit, we have the arrangement of *Figure 14.16*, which is the most satisfactory layout. Independence of frequency on the HT supply is secured by adjusting the potentiometer, as before, and in virtue of the rearrangement of the oscillator section the screen potential is now fixed and serves to isolate the oscillator circuit from happenings at the valve anode.

Summarizing the findings on LC oscillators, the simple Hartley and Colpitt circuits are characterized by: large voltage outputs at good waveform, and—by making the valve a beam tetrode or other power valve—large output powers. Automatic amplitude stabilization is simply obtained, but only poor stabilization of amplitude and frequency against variations of load. It is

LC OSCILLATORS

difficult to vary the frequency over a wide range without severely influencing the output.

When these simple circuits are followed by a buffer valve, the frequency is substantially independent of the load. The output voltage is generally smaller if the buffer valve load is a resistance but not if it is a transformer or another tuned circuit. The output power depends on the buffer valve. There is the cost of the extra valve. The purity of the output waveform is good.

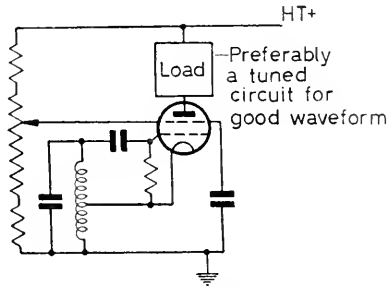


Figure 14.16

With the electron coupled oscillator the frequency is independent of the load and may be made independent of the HT voltage, over a range. The output voltage may be large, and also the output power if the valve is a beam tetrode or power pentode. The purity of the output waveform is poor, and requires resuscitating with a further tuned circuit.

Danger—red hot screens

In designing Hartley or Colpitts oscillators to deliver appreciable powers—of the order of watts and above—there are good reasons for employing beam tetrodes or power pentodes because of their greater efficiency compared with triodes. If this is done it must be borne in mind that when the anode swings negative, the anode potential will be extremely low—possibly even lower than the cathode—and that this state of affairs coincides with maximum positivity at the grid and hence maximum cathode current. Under these circumstances nearly all the cathode current will flow to the screen, an electrode normally required to pass only a fraction of the cathode current. It is necessary to pay special attention to the maximum allowable screen dissipation for the valve, and to check the product of average screen current—as measured by a moving coil instrument—and screen potential employed. Failure to remember this, and the use of a screen potential which might be quite suitable when the valve is used as an amplifier, are liable to lead to oscillators which run with their screens at red heat. This promotes the liberation of adsorbed gases, the bombardment of the cathode by gas ions, and the early demise of the valve.

Applications of LC oscillators of interest to the electrobiologist are cathode ray tube beam brightening¹, tape recorders², multi-channel oscillography³ and non-lethal high voltage supply⁴.

Blocking oscillator

This is a variety of LC oscillator which is sometimes useful. Though they are not continuous sine wave oscillators this seems an appropriate

SINE WAVE OSCILLATORS

point at which to introduce them. They are a means of generating an exceedingly brief pulse, say, less than 1 microsecond, or they may be used to generate approximately triangular waves. Suppose we take an ordinary LC oscillator of the type first discussed and transpose the pick-up coil and the stabilizing RC network, as shown in *Figure 14.17*. This has no effect on the operation of

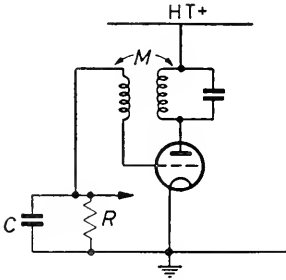


Figure 14.17

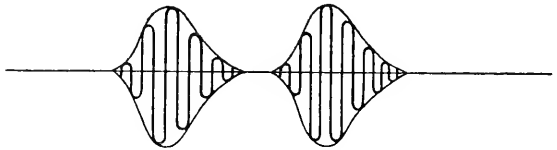


Figure 14.18

the circuit. Now suppose we increase CR . The time taken for the bias to adjust itself to a small reduction in the loop gain, perhaps caused by a small decrease in the HT supply, becomes longer till a stage is reached at which the oscillations apparently die away spontaneously. If now, to counteract this, we increase the loop gain, usually by increasing M , we find the amplitudes of oscillations become themselves oscillatory. Oscillations die away as before, then the bias leaks away, whereupon fierce oscillation becomes once more possible. Oscillations rapidly build up, generating a bias which eventually rises to a level too high to sustain them, then they die away once more (*Figure 14.18*). This state of affairs is called 'squegging' and is the halfway house to the production of a blocking oscillator. To achieve the latter we go on increasing the loop gain, and the effect of this is to cause the oscillations to grow so fast that the capacitance is charged extremely quickly. Eventually a state is reached where the valve is cut off after only half a cycle of oscillation: this is the blocking oscillator. In that half-cycle enough grid current flows to charge the capacitor to a bias voltage which will prevent a further cycle of oscillation. The circuit is quiescent until this bias voltage has leaked away sufficiently for oscillation to recommence. The waveform across R and C

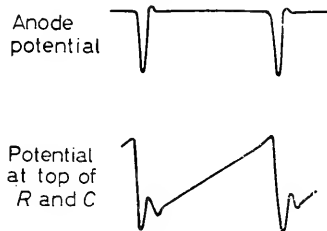


Figure 14.19

is therefore quasi-triangular. At the anode a train of negative-going pips is available, of duration approximately half a cycle of the frequency at which the tuned circuit is resonant (*Figure 14.19*).

CRYSTAL OSCILLATORS

CRYSTAL OSCILLATORS

In applications where extreme stability of frequency is required, such as in the production of time markers, quartz crystal oscillators are used; for with these, provided the crystal is of good quality, the circuit must of necessity oscillate at the correct frequency if it is to oscillate at all.

Oscillation secured with the aid of a quartz crystal is associated with a *mechanical* resonance, which, in virtue of the mechano electric coupling by the Piezo effect, appears between the crystal terminals as an electrical resonance of remarkable sharpness. The electrical equivalent circuit has the form of *Figure 14.20*, from which it is clear that two forms of resonance are possible,

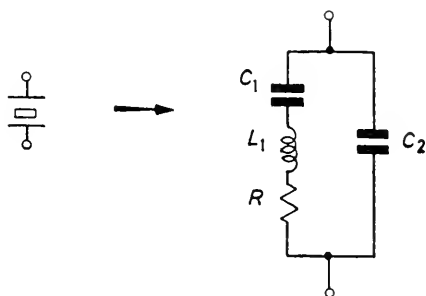


Figure 14.20

one series and one parallel. In crystal oscillators we are concerned only with the parallel resonance. Quartz is an almost perfectly elastic material which implies that if shocked mechanically into vibration the vibrations will die away extremely slowly. We have seen that in LCR circuits this is a state of affairs associated with high Q , and it is to be expected that the effective Q of a crystal will be large, and so it is.

A certain 450 kilocycles crystal had an equivalent L of 12.8 henries, $C_1 = 0.01$ pF, R of 1,600 ohms and C_2 of 18.3 pF. In computing the Q value we can take $C = 0.01$ pF, since the value of capacitance formed by C_1 and C_2 in series is almost the same as C_1 alone. Then

$$Q = \left(\frac{L}{C}\right)^{1/2} \cdot \frac{1}{R} = \left(\frac{12.8}{10^{-14}}\right)^{1/2} \cdot \frac{1}{1600} = 22,400$$

Similarly a 3 Mc/s crystal described by Scroggie* had $L = 0.127$ H, $C_1 = 0.022$ pF, $C_2 = 8$ pF and $R = 30$ ohms, and

$$Q = \left(\frac{0.127}{2.2 \times 10^{-14}}\right)^{1/2} \times \frac{1}{30} = 80,000$$

There are two classical crystal oscillator circuits, the Pierce and the 'TATG'.

Pierce oscillator

The Pierce oscillator is extremely simple and has the appearance of *Figure 14.21*. To see how it works we have to remember the stray capacitances

* *Radio Laboratory Handbook.*

SINE WAVE OSCILLATORS

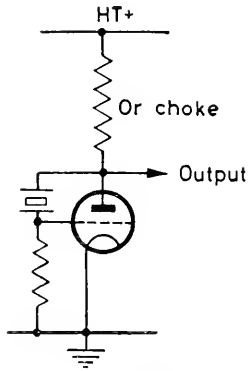


Figure 14.21

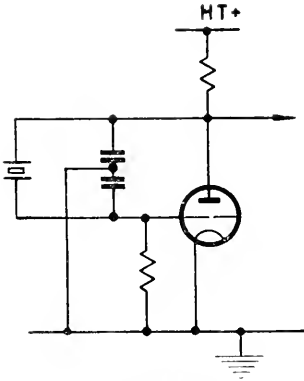


Figure 14.22

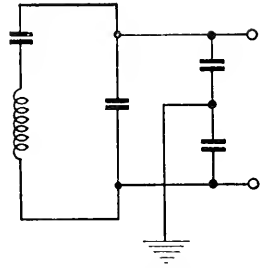
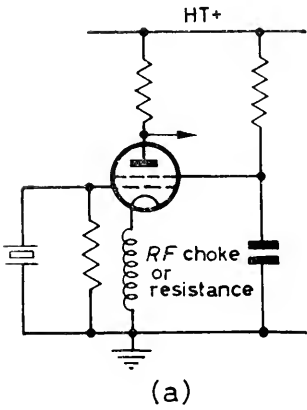
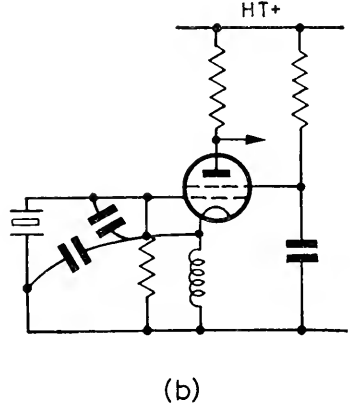


Figure 14.23



(a)



(b)

Figure 14.24

CRYSTAL OSCILLATORS

between grid and earth, and between anode and earth. Putting these in (*Figure 14.22*) and remembering that the crystal is to act as a parallel tuned circuit, we see that we have a modified Colpitts oscillator in which some of the tuning capacitance is direct across the coil and some of it is split, Colpitts-wise, to effect the virtual tap on the tuned circuit (*Figure 14.23*).

The output from a Pierce oscillator can be taken out directly to a high-impedance load such as another valve as shown, or by electron coupling, in which case the oscillator circuit undergoes the same kind of transformation that occurred between *Figure 14.6* and *14.12*. The circuit as published will have the appearance of *Figure 14.24a*, but a truer representation is *Figure 14.24b*. The purpose of the choke inductance is to carry the cathode current while allowing the cathode to be an active electrode driven by the stray capacitances as shown. As usual with electron coupled schemes the purity of the output is poor and the anode load is often made a tuned circuit, where this matters.

TATG oscillator

This has the form of *Figure 14.25*, from which it will be seen that an LC tuned circuit is required as well as the crystal, but the circuit has the merit of giving rather greater output. The mode of operation is complicated; the name arises from an equivalent LC oscillator which is of little interest to electrobiologists and has not been mentioned, the tuned anode, tuned grid (*Figure 14.26*). The feedback is via the valve grid-anode capacitance, and can

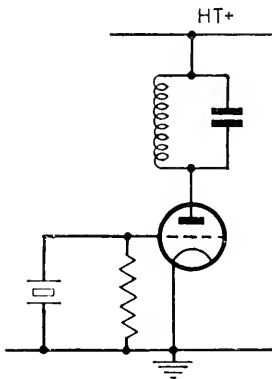


Figure 14.25

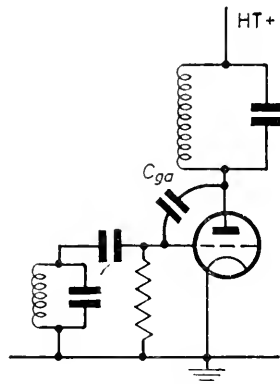


Figure 14.26

be shown to be positive at a frequency just below that at which the tuned circuits are resonant. In the crystal oscillator, setting-up procedure involves monitoring the anode currents as the tuned circuit is adjusted. As adjustment proceeds it will be found that the anode current passes through a minimum having one steep and one gradual side. The tuned circuit should be left on the gradual side of the minimum as shown in *Figure 14.27*. The oscillations are maximal when the anode current is adjusted to the bottom of the trough, but this is an unstable point, since any upward drift in the tuned circuit resonant frequency will cause the circuit suddenly to cease oscillating.

SINE WAVE OSCILLATORS

Modern crystal oscillators

Crystal oscillators of the type so far discussed are primarily intended for radio work and the frequencies of crystals suitable for such circuits are rather high to be of direct use as timers in electrophysiological circuits. Thus it was at one time usual to employ 100 kc/s crystal oscillator and 'count

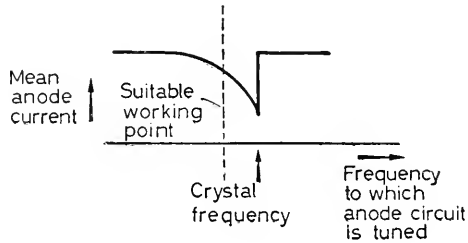


Figure 14.27

down' 10 times to produce 10 kc/s (100 μ sec marks), count down 10 times again to produce 1 kc/s (1 msec mark) and so on. Since marks are seldom required at 10 μ sec intervals, it is wasteful to produce an unnecessarily high frequency such as 100 kc/s and then have to cut down 10 times before producing a useful frequency. Modern four-terminal crystals such as the Standard Telephone and Cable 4023 can be had which oscillate at frequencies as low as 4 kc/s. The circuit recommended by Standard Telephones for crystal oscillators of this type is given in Figure 14.28. The feedback path is

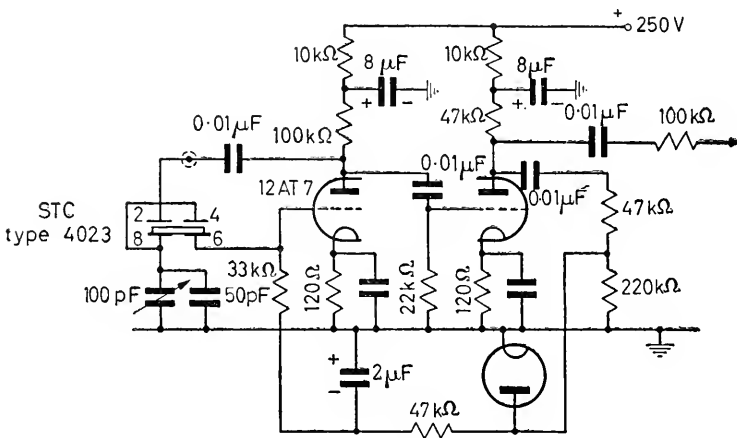


Figure 14.28

mechanical through the crystal itself. The circuit has the D'Argumbeau amplitude regulating device to be discussed under RC oscillators.

Timers for electrophysiological work employing crystal oscillators have been described by Dickinson⁵ (TATG), Kay⁶ (Modern) and Attew⁷ (electron coupled Pierce).

RC OSCILLATORS

Danger—cracked crystals

It must always be remembered that the resonance of the crystal is mechanical and, since quartz is a brittle material, that if the oscillations are allowed to become excessive the crystal will break. There is no need for this, since the function of a crystal oscillator is primarily to give the correct frequency, not to supply large powers; power can always be had by appropriate subsequent valve amplification. In developing crystal oscillators the HT supply should initially be low, and the a.c. voltage across the crystal monitored to ensure that it does not exceed that recommended by the manufacturers.

RC OSCILLATORS

The RC oscillators are closely related to the RC tuned acceptor amplifiers employing positive feedback; virtually they are tuned amplifiers with the input terminal arrangement left out and AB made equal to one.

It is therefore to be expected that there would be two of them, a 'phase shift' oscillator and one involving the RC band-pass filter, $a = 1$. In fact nearly all RC oscillators are of one of these two kinds, though the latter type are known rather indiscriminately as Wien bridge oscillators. The author feels that this is a bad name, since the RC band-pass filter is only half of a Wien bridge, and is an acceptor network, whereas the Wien bridge is a null-transmission rejector network. Only in some circuits, such as *Figure 14.35*, is there an actual Wien bridge. However, the name is brief and seems to be generally accepted.

RC oscillators are nowadays almost universal for the generation of test oscillations in the frequency range 100 kc/s–1 c/s or even lower, their advantage being the cheapness with which such wide ranges may be covered. The method is to have either R (or C) variable in switched steps, and C (or R) continuously adjustable. If the lowest frequency to which the oscillator is to tune is of the order of 20 cycles, design usually proves easier if a variable C is used—a radio type multi-gang tuning capacitor—with switched R . If the apparatus is to go down to a much lower frequency, then the R values required become undesirably large since the capacitors usually have a maximum value of some 500 pF per section; if the R values are large the frequency determining elements in the circuit are prone to interference from stray electric fields and elaborate screening becomes necessary; also there may be transgressions of the valve maker's rules about the maximum value of grid-cathode resistance.

Thus for very low frequency RC oscillation it is better to have C switched in steps and R continuously variable. If this is done the ganged variable resistances must be of exceptionally good quality. The ganging of ordinary double potentiometers is not as accurate as that of ganged capacitors and when the frequency of an oscillator employing one is altered it will be found that the automatic level control device is kept hard at work.

In outline the phase shift oscillator is shown in *Figure 14.29a* and the Wien bridge type in *Figure 14.29b*. For a fixed frequency oscillator the phase shift version is preferable on the ground of simplicity, but the Wien bridge type is more usual in wide-band oscillators because fewer frequency determining components are required. The Wien bridge type requires an overall

SINE WAVE OSCILLATORS

amplifier gain of 3 and oscillates at $\omega = 1/RC$. The phase shift oscillator requires a gain of about 8 and oscillates at $1/(3)^{1/2}RC$ if a is about 10. If $a = 1$, the gain required is 29 and $\omega = 1/(6)^{1/2}RC$.

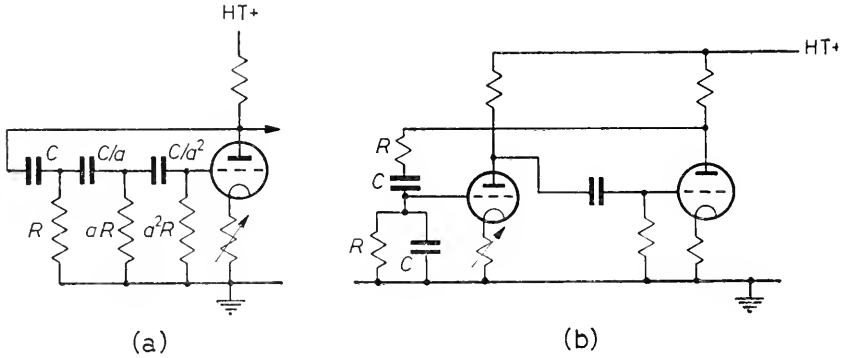


Figure 14.29

Arrangements for maintaining the level of oscillation have to be more elaborate than the simple self-biasing device which is adequate for LC oscillators. This is because the RC frequency-determining networks have *of themselves* only a very small Q value; one could not possibly tolerate the pulsatory currents which are permissible in LC oscillators because there is no device with a Q of several hundred to re-create a pure sine waveform. Thus the valves must amplify in class A, without distortion, and some other means must be found of regulating the gain.

D'Arquimbeau's method—There is a variety of pentode known as 'variable μ ' (Figure 14.30). As a result of grading the mesh size of the control grid along the axis of the valve it can be shown that the amplification obtainable from the valve is a function of the mean grid bias (Figure 14.31). Clearly if the



Figure 14.30

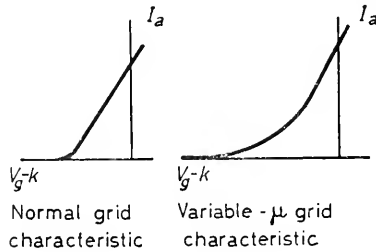


Figure 14.31

input signal is large enough to sweep over an appreciable proportion of this curve, the output of a variable μ stage is distorted. However, if the signal is small enough only a short piece of the curve is used and this may be regarded as being straight. Use of this fact may be made in the Wien bridge oscillator. In Figure 14.32, part of the output is tapped off and shunt rectified by the small diode in the second valve to produce a negative bias which controls the gain of the first valve. The capacitance of C_2 has to be sufficiently large for its reactance to be negligible compared with the frequency determining R 's and

RC OSCILLATORS

C 's at all times. R_2 and C_2 form a low-pass filter to smooth the control bias.

The system works well and the output waveform is quite pure provided the level of oscillation is held down to a volt or two. The level is controlled by the setting of the potentiometer.

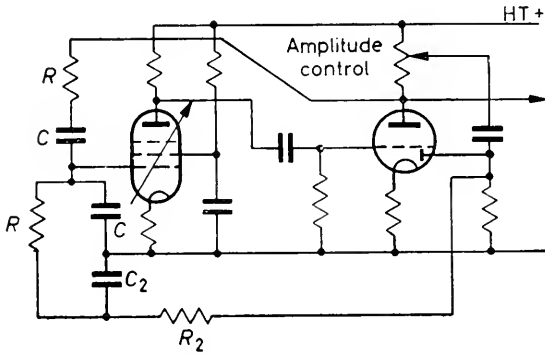


Figure 14.32

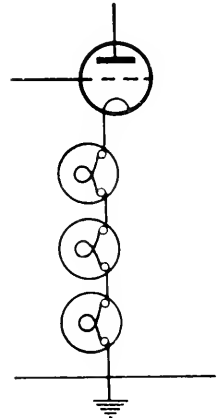


Figure 14.33

Lamp method—A very much simpler scheme makes use of the temperature coefficient of resistance of metals. The method is equally applicable to phase shift and Wien bridge oscillators. If the cathode resistance of one of the valves is composed of one or more metal filament electric bulbs (Figure 14.33), then, as the valve is in class A, when the amplitude of the input signal increases the average cathode current is unaffected but the R.M.S. cathode current rises. Hence the temperature of the bulb filaments also rise, and so does their resistance. This increases the negative feedback applied to the valve and reduces the loop gain, counteracting the original increase in input.

For the scheme to be effective the bulb filament temperature must be a fast-changing function of the R.M.S. current, which implies that the rated bulb filament current must not be too greatly in excess of the average valve anode current. The latter is typically 5 mA. It is possible to get rear-light bulbs for bicycles which consume only 40 mA; these are very suitable.

Thermistor method—A popular modern method of stabilizing the level of oscillation is to use a thermistor. It consists of a small piece (about the size of a pinhead) of a material having a very large negative temperature coefficient of resistance. Their symbol is Figure 14.34 and they may be used in various ways. In Figure 14.35, C' is large enough to have negligible reactance even at the lowest frequency of oscillation. The feedback is positively to the first valve grid, in the normal manner for Wien bridge oscillators, and also negatively to the first valve cathode via the thermistor. When the amplitude of the oscillations rises, the thermistor heats up and its resistance falls, increasing the negative feedback.

A very neat circuit (Figure 14.36) was shown to the author by K. E. Machin. A long-tailed pair is used, so that the feedback to the frequency determining network is negative from the left-hand anode and positive from the right-hand. The thermistor is across the output and heats up if the latter

SINE WAVE OSCILLATORS

tends to increase, increasing the negative feedback. The outputs are in push-pull and the magnitude determined by the variable resistor. It is quite easy to get 50 V peak to peak with good waveform from each anode, with this circuit.

A complete oscillator, using the Wien bridge principle, and very suitable for checking electrophysiological apparatus, has been published by Scroggie⁸,

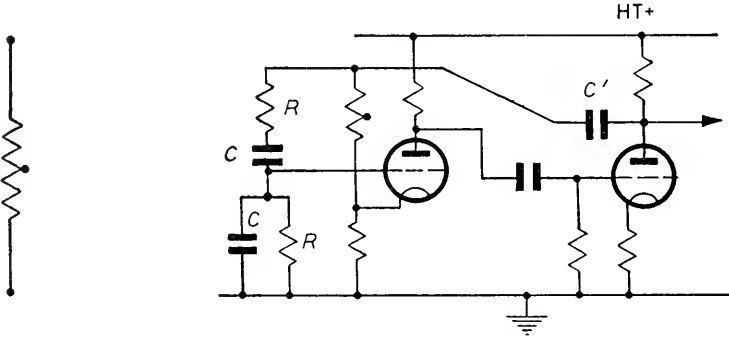


Figure 14.34

Figure 14.35

see also Sinfield⁹. A very ingenious form of phase shift oscillator of novel design has been described by Raistrick¹⁰. It is claimed that no regulating device is required, and only one resistance has to be altered to adjust the frequency; three valves are, however, necessary. A similar device which can

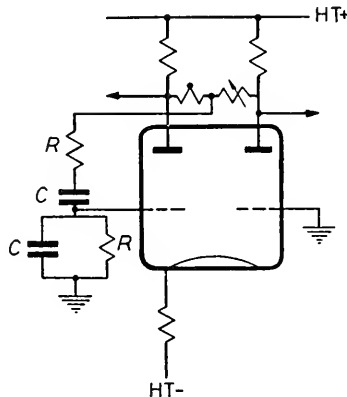


Figure 14.36

be used as an oscillator or as a tuned amplifier has been published by Morton¹¹.

REFERENCES

- ¹ ATTREE V. H. *J. Sci. Instrum.* 26 (1949) 257
- ² CARTER J. M. Designing a tape recorder *Wireless World* 59 (1953) 164
- ³ DONALDSON P. E. K. A multiple-channel oscilloscope for electrophysiology *Electron. Enging.* 39 (1957) 78
- ⁴ MACLEOD K. C. *J. Television Soc.* 5 (1949) 201

REFERENCES

- ⁵ DICKINSON, C. J. Electrophysiological technique *Electron. Enging. Monog.* p. 98
- ⁶ KAY, R. H. A time-marker for electrophysiology *Electron. Enging.* 28 (1956) 452
- ⁷ ATTEW, J. E. Decade multi-vibrator design *Wireless World* 58 (1952) 114
- ⁸ SCROGGIE, M. G. Audio signal generator *Wireless World* 55 (1949) 331
- ⁹ SINFIELD, L. F. Extended range i.f. oscillator *Wireless World* 60 (1954) 596
- ¹⁰ RAISTRICK, W. G. Phase-shift oscillators *Wireless World* 59 (1953) 409
- ¹¹ MORTON, A. Q. Oscillator filter unit *Wireless World* 59 (1953) 129

SQUARE WAVE OSCILLATORS

A square wave is one having the form of *Figure 15.1*. The ratio a/b is called the mark-space ratio. Square waves are useful for testing the transient

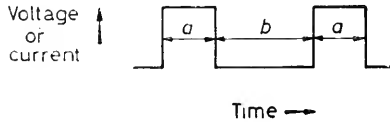


Figure 15.1

response of recording equipment, and square wave oscillators are complementary in function to sine wave oscillators, which check the steady-state response.

The 'squared-up' sine wave

Square waves may be made from sine waves by amplification and clipping, and many commercial, primarily sine wave, test oscillators improve their sphere of usefulness by having a square wave outlet obtained on this principle. The procedure is merely to feed a valve with a sine wave input which is, as it were, much too big for it (*Figure 15.2*). When the input goes positive the

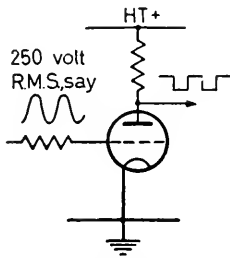


Figure 15.2

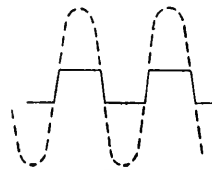


Figure 15.3

grid cannot follow it beyond a certain point because it is 'caught' by diode action between it and the cathode. When the input goes negative the anode cannot follow it beyond a certain point—the HT plus potential—because the valve is cut off. The anode waveform is therefore moderately square in that the tops and bottoms of the waves will be quite flat, but the sides will still be rather sloping (*Figure 15.3*). However, if the process be repeated in a further valve the final product is likely to be square enough for most purposes.

In pieces of apparatus where square waves only are required it is a waste of effort to make sine waves and then spoil them again, and we use a circuit

THE MULTI-VIBRATOR

specifically for the production of square waves*. This circuit is the multi-vibrator; it has the additional advantage that the mark-space ratio is variable, whereas in the case of the squared-up sine wave it is necessarily about unity.

THE MULTI-VIBRATOR

The multi-vibrator exists in two forms, the symmetrical and the cathode coupled.

Symmetrical multi-vibrators

The symmetrical version is shown in *Figure 15.4*. The circuit oscillates because the feedback is clearly positive, and the degree of feedback is much higher than that required just to maintain oscillations. Thus the level of

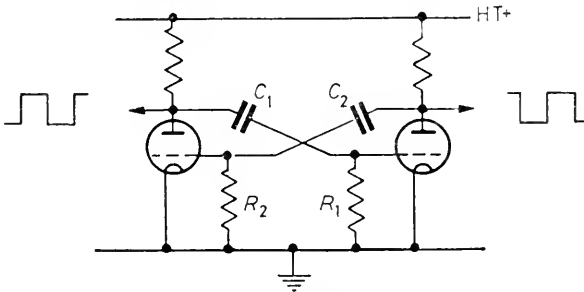


Figure 15.4

oscillation is limited only by the maximum excursions of which the valve anodes are capable, i.e. from HT positive (valve cut off) to a level of anode potential corresponding to zero, or slightly positive, grid bias. We shall call this latter state 'conducting hard'.

This circuit and those which we shall consider in this and the next chapter are all non-linear ones in which the anode potential of a valve is not a replica of the grid potential, as it has been in most of the valve circuits we have dealt with up to now. With the non-linear or 'pulse' circuits we often find valves being used in the rather simple role of switches. They are either 'off' (or 'cut off') or 'on' (or 'conducting hard').

Design procedure for pulse circuits may be carried out along the same general lines as for amplifier circuits. That is to say, a working region is constructed on the valve anode characteristic and a load line drawn. Because non-linear operation is intended the working region extends to the anode voltage axis and somewhat outside the characteristic for zero grid bias. 'Off' and 'on' are represented by the blobs on the load line (*Figure 15.5*).

High speed of operation on transients in pulse circuits is secured by, amongst other things, the use of anode loads of rather low resistance, of the order 5–10 k Ω instead of 50–100 k Ω . This has the effect of making the load line steep, and in order to secure a swing of anode voltage comparable with that of *Figure 15.5* a high anode current and anode dissipation become

* Except in cases where the frequency has to be very stable, when it is best to square up the output of a crystal oscillator.

SQUARE WAVE OSCILLATORS

necessary. The question that then arises is to what extent the boundaries of the working region may be transgressed. For example, under what circumstance is a 'conducting hard' point like *Figure 15.6* permissible, where both the anode dissipation and the maximum cathode current are being exceeded?

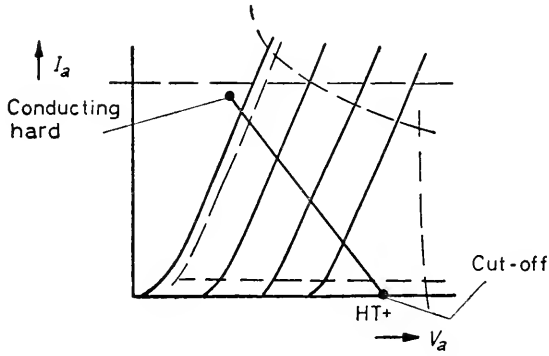


Figure 15.5

The answer is that the anode dissipation may be instantaneously exceeded, provided it is not exceeded on the average, the average being taken over a period short compared with the thermal time-constant of the anode structure. The maximum cathode current rating can apparently also be exceeded instantaneously provided it is observed on the average. What is happening

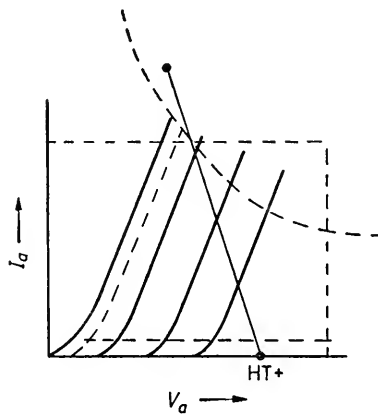


Figure 15.6

here is that the space charge is being depleted to supply the anode current, a practice which should be indulged in moderately; it should be remembered that the space charge gives the cathode a measure of protection from ionic bombardment should the valve possess some residual gas. It is probably wise to consult the valve makers in cases where appreciable instantaneous excess cathode currents are contemplated.

THE MULTI-VIBRATOR

In the symmetrical multi-vibrator the circuit alters very rapidly between the states V_1 conducting, V_2 cut off, and V_2 conducting, V_1 cut off. The times for which it dwells in each of these states depend on the time constants C_1R_1 and C_2R_2 respectively. Let us suppose V_1 has just been switched on and V_2 off. V_1 anode potential has fallen to its low 'conducting hard' value and has taken, via C_1 , V_2 grid to a potential far below cut off (A in *Figure 15.7*). Current will flow into the right-hand plate of C_1 , via R_1 , causing it

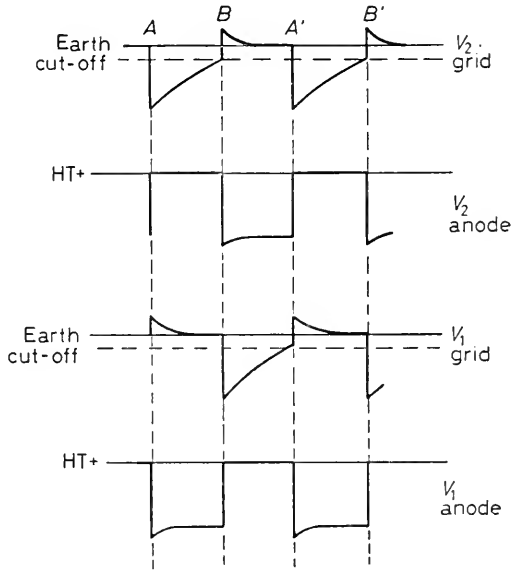


Figure 15.7

and V_2 grid to return exponentially towards earth. When V_2 grid has reached cut off, anode current begins to flow in V_2 and V_2 anode begins to fall in potential. This carries V_1 grid negative, reducing V_1 anode current and begins to raise V_1 anode. This rise is communicated to V_2 grid so that a cumulative action is initiated as a result of which V_1 is completely cut off and V_2 is switched to hard conduction (B in *Figure 15.7*). Notice that V_2 grid is carried slightly positive at this instant. Grid current then flows into V_2 from C_1 , re-establishing the potential difference across it at its value just before A .

The other half of the cycle takes place in similar fashion, the roles of V_1 and V_2 being reversed (B to A' in *Figure 15.7*).

Now the length of the period AB depends on the time it takes V_2 grid to rise to cut-off bias, which in turn depends on the time constant R_1C_1 and on how far negative V_2 was originally carried. The latter quantity depends on the magnitude of the negative swing at V_1 anode, which is strongly dependent on the HT and the parameters of V_1 . Similarly BA' depends on R_2 and C_2 , the HT and the parameters of V_2 . From this it should be clear that whilst with this circuit the mark-space ratio can easily be varied by differential variation of the products R_1C_1 and R_2C_2 , the frequency of oscillation is not a

SQUARE WAVE OSCILLATORS

property of the (relatively stable) passive R 's and C 's as it was with the RC sine wave oscillators, but is also dependent on the HT—which can be stabilized, though it is a nuisance to have to do so—and the valve parameters, liable to progressive alteration as the valves age. For this reason the multi-vibrator is not used as a source of known and stable frequency, it being better where frequency is important to square up the output of a sine wave oscillator, as mentioned earlier.

Pentode multi-vibrator

The rapidity of transition between the two states depends on the grids having low capacitances to earth, and being driven from low source resistances (*Figure 15.8*). The capacitances can be greatly reduced by using

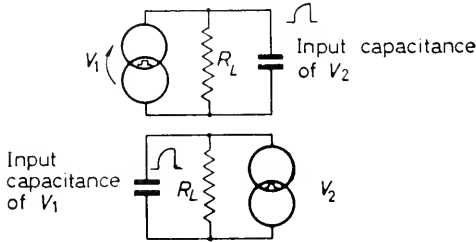


Figure 15.8

pentodes instead of triodes, and if pentodes are used the source impedances are substantially the anode loads. The circuit has the appearance of *Figure 15.9*. The screen decoupling time constants must be long compared with the time of a cycle of oscillation. If this is not possible because the latter is

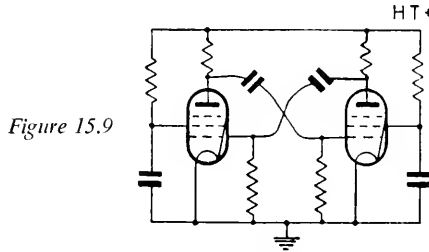


Figure 15.9

itself low, the screens will require stabilization, usually, by a soft diode. The time constant of the transitions of the circuit, t in *Figure 15.10*, will

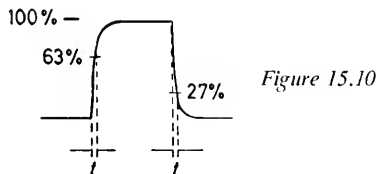


Figure 15.10

be R_L times the total shunt capacitance across each grid, including strays. If it can be assumed that the bulk of this is composed of C_{gk} , then for a

THE MULTI-VIBRATOR

given gain, $g_m R_L$, we see that a figure of merit for pentodes for this and other types of high speed pulse work is merely g_m / C_{gk} .

Screen-coupled multi-vibrators—Sometimes it is desirable to make a multi-vibrator work into an inductive load such as a relay coil or lightly loaded transformer. If a triode multi-vibrator is used the presence of the inductance in place of pure resistive loads seriously modifies the mode of operation, and renders it highly dependent on the nature of the inductive load. A solution is to use pentodes and obtain the feedback signals from the screens. The anodes are then free for the driving of almost any kind of load, without affecting the operation of the multi-vibrator proper (*Figure 15.11*). It must

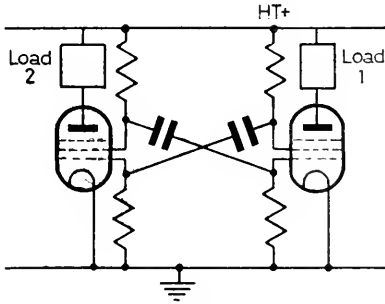


Figure 15.11

be borne in mind that so far as speed of transition is concerned this is a triode, not a pentode, multi-vibrator, as there is now considerable Miller capacitance from screens to control grids.

Cathode coupled multi-vibrators

Consider a circuit like *Figure 15.12*. V_1 is a cathode follower and V_2 a triode amplifier. If V_1 grid is made very negative with respect to earth, V_1 is cut off and V_2 acts as an amplifier valve automatically biased by R_2 to some anode current such that the anode potential is, say, V_A .

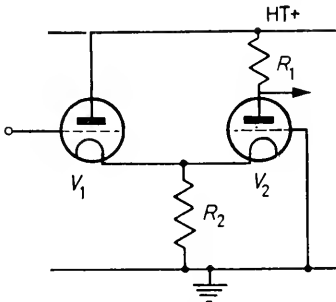


Figure 15.12

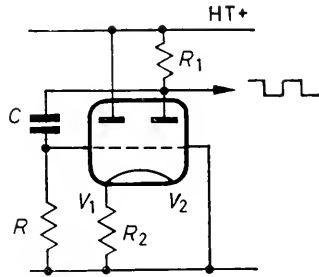


Figure 15.13

Now let V_1 grid be made positive with respect to earth, perhaps 50 V or so. V_1 cathode will follow it, taking V_2 cathode positive as well, cutting off V_2 , so that the anode potential of V_2 rises to HT positive.

SQUARE WAVE OSCILLATORS

Let the least negative potential at which V_1 is cut off, and the least positive potential at which V_2 is cut off, be called V_B and V_C respectively. Then when V_1 grid is between V_B and V_C the circuit will behave as an approximately linear amplifier, and if a feedback path be established between V_2 anode and V_1 grid, the feedback is clearly positive.

In the cathode coupled multi-vibrator the feedback is via a capacitance C , a leak R being provided from V_1 grid to earth (*Figure 15.13*). Circuits of this kind can operate under either of two regimes, according to whether or not V_1 passes significant amounts of grid current. The frequency is determined largely by the product CR ; the mark-space ratio by the regime.

Regime 1—no grid current—In this case the circuit performs oscillations with a mark-space ratio of approximately unity, the waveforms having the form of *Figure 15.14*. The amplitude of the output is $(HT+) - V_A$, that is,

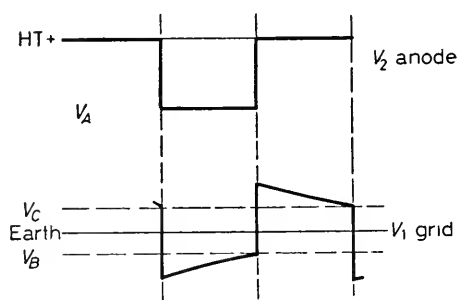


Figure 15.14

the voltage drop across the anode load when V_2 is conducting hard. We have for V_2 ,

$$\begin{aligned} \text{Anode current} &\propto \frac{1}{\text{bias}} \text{ approximately} \\ &\propto \frac{1}{\text{anode current} \times R_2} \end{aligned}$$

and so
$$\text{anode current} \propto \frac{1}{(R_2)^{1/2}}$$

$$\text{Multi-vibrator output} = \text{Voltage drop in } R_1 = \text{Anode current} \times R_1$$

$$\propto \frac{R_1}{(R_2)^{1/2}} \text{ roughly}$$

If we try—by reducing R_2 —to increase the amplitude of the output beyond a certain point, the circuit moves into the other regime and the mark-space ratio changes rapidly. The output has the appearance of *Figure 15.15*. The explanation is as follows: for satisfactory cathode follower action to occur in V_1 , V_1 cathode must always be able to supply enough current to keep its own potential positive with respect to V_1 grid. When R_2 is reduced the current which the cathode has to be able to emit to follow a *given* positive grid potential has also to increase. As we reduce R_2 a point is reached at

THE MULTI-VIBRATOR

which the valve is unable to do this, and the cathode is unable to maintain itself positive with respect to the grid on the extreme positive excursions of the latter. When this happens we have:

Regime 2—significant grid current flow—The effect of the grid current is to give C a mean charge tending to bias V_1 grid negatively. The entire grid waveform is depressed, as shown in *Figure 15.16*, producing the change in

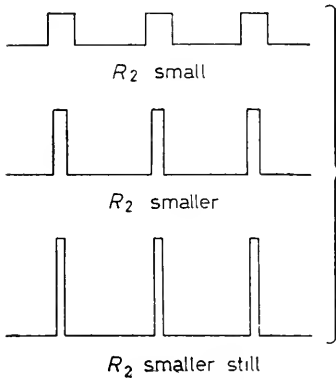


Figure 15.15

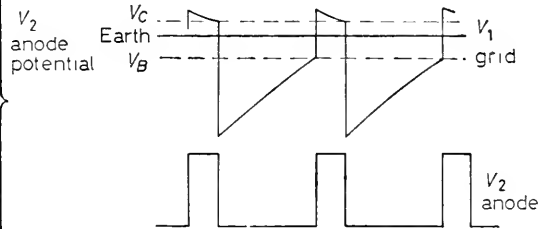


Figure 15.16

mark-space ratio alluded to. It is not possible to produce a simple expression for the value of R_2 at which the change of regime takes place. In a cathode coupled multi-vibrator of the author's, using an ECC 81 valve and an anode load of 50 k Ω the critical value of R_2 was about 3 k Ω . Above this the mark-space ratio remained at about unity and the amplitude of the output fell. Below this the output rose and the mark-space ratio altered rapidly, as in *Figure 15.15*. As a general rule with pulse circuits the quickest way to produce a workable design is by experiment, it having been first established that the working points of the valves are—at least on the average—safely within their working regions.

In conclusion, then, the cathode coupled multi-vibrator is characterized by having fewer components than the symmetrical type, and the frequency may be varied—whilst preserving the mark-space ratio—by altering only one RC product. The mark-space ratio is determined by the ratio of anode load and common cathode resistance, and so is the amplitude. Frequency, mark-space and amplitude are all affected by the HT and valves used.

Application of multi-vibrators

Although the multi-vibrator is of no use as a frequency standard itself, it can easily be synchronized to run at the same frequency as, or—more important—at a sub-harmonic frequency of a master frequency standard such as a crystal oscillator. Use is made of this fact in some kinds of timer, where the crystal oscillator runs at 100 kc/s, say, and controls the frequency of multi-vibrator No. 1, running at 100 kc/s. Multi-vibrator No. 1 controls multi-vibrator No. 2, running at 1 kc/s, and so on. The process is known as 'counting down'.

SQUARE WAVE OSCILLATORS

Counting down is achieved in the symmetrical multi-vibrator by injecting positive 'pips' of voltage at the control or master frequency into one of the grid circuits. If the control pips are at 100 kc/s and the multi-vibrator is to count down 10 times, then it is arranged that, in the absence of any synchronization, the circuit would oscillate rather below 10 kc/s. The tenth synchronizing pip then comes along and trips the multi-vibrator slightly early (Figure 15.17—cf. Figure 15.7). It is clear from Figure 15.17 that the amount

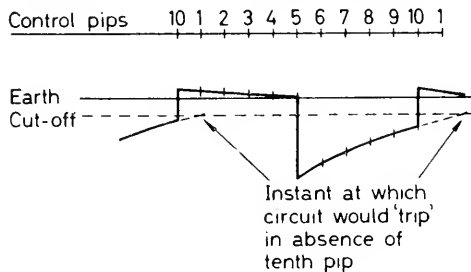


Figure 15.17

of control pip injected must be carefully watched, otherwise the oscillator will be tripped on the ninth if the pips get too big, and fail to be tripped by the tenth if the pips get too small.

Circuits of electrobiological interest which employ multi-vibrators have been described by Dickinson¹ (timer employing symmetrical multi-vibrators), Attew² (timer employing a variety of cathode coupled multi-vibrators) and Attree³ (device for producing stabilized heater supplies, employing screen-coupled multi-vibrators).

REFERENCES

- ¹ DICKINSON, C. J. Electronic engineering monograph *Electrophysiological Technique* p. 98
- ² ATTEW, J. E. Decade multi-vibrator design *Wireless World* 58 (1952) 114
- ³ ATTREE, V. H. A low voltage stabilizer with saturated diode control *Electron. Engng.* 25 (1953) 71

TRIGGERED PULSE GENERATORS

Triggered pulse generators are circuits which are inactive until the arrival of a brief control or 'trigger' pulse, when they go through a cycle of activity during which a special waveform is generated and at the end of which they are 're-cocked' and ready to repeat the performance upon the arrival of



Figure 16.1

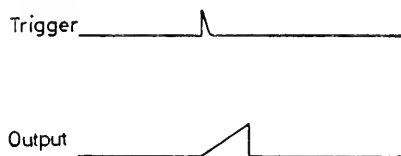


Figure 16.2

another pulse. The most important waveforms required of circuits of this kind are a single square wave (Figure 16.1) or a single triangular wave (Figure 16.2).

SINGLE SQUARE WAVE GENERATORS

The cathode coupled flip-flop

This circuit has the appearance of Figure 16.3, from which it is evident that it is closely related to the cathode coupled multi-vibrator; in fact it may be regarded as a multi-vibrator in which V_2 grid has been given so much positive

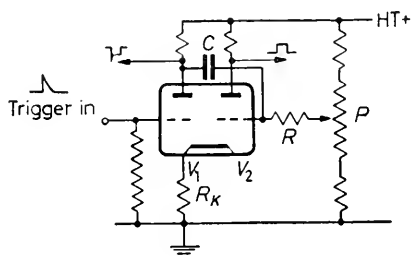


Figure 16.3

bias that in the absence of an outside stimulus oscillation is prevented. In the 'cocked' condition V_2 is conducting heavily because of the large positive bias on V_2 grid, and a considerable current flows down R_K , making the common cathode sufficiently positive to cut off V_1 .

The circuit is 'fired' by applying to V_1 grid a short positive pulse which allows V_1 to conduct. V_1 anode potential then falls, driving V_2 grid negative

TRIGGERED PULSE GENERATORS

via C , reducing the cathode potential and hence tending further to increase the current in V_1 . The result is the now-familiar cumulative action, as a result of which V_2 becomes cut off and V_1 heavily conductive.

Current now flows via R into the right-hand plate of C , causing V_2 grid to rise exponentially towards the potential of the slider of P . At length, conduction begins again in V_2 , initiating a cumulative action in the reverse direction, re-cocking the circuit in its original state. Waveforms are sketched in *Figure 16.4*. An output may be taken from either anode according to

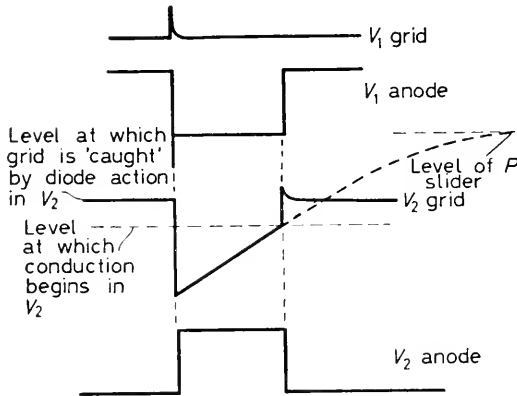


Figure 16.4

the polarity required. The duration of the pulse may be controlled by varying R , C , or the slider of P . The amplitude of the pulse depends on the valve, the HT, and is related to the ratio of the relevant anode load to R_K . The pulse duration also depends on these three quantities. A high rate of rise

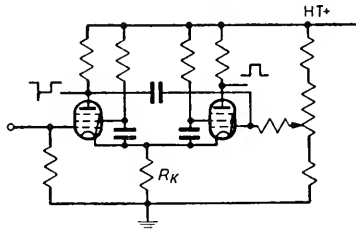


Figure 16.5

and fall at the beginning and end of the pulse is secured by having the anode loads low and by minimizing shunt capacitances, particularly by using pentodes (*Figure 16.5*). Remember that the screens have to be decoupled to the *cathode*, since pentode action depends on stabilizing the screen cathode potential difference.

Triggered transitron

Just as the flip-flop is a biased cathode coupled multi-vibrator, the triggered transitron is a biased version of the transitron oscillator. The latter does

SINGLE SQUARE WAVE GENERATORS

not seem to the author to be electrobiologically interesting and so has not been described, but the triggered version is important and will be dealt with now.

In the transitron a pentode valve is used, but not in the classical pentode manner, where the screen-cathode and suppressor-cathode potential differences are constant. Instead, the screen is used as a kind of subsidiary anode and the suppressor as a second control grid. Not many pentodes make good transitrons because the controlling effect of the suppressor of most pentodes is insufficient. At the time of writing the Mullard EF 50 and the Mazda 6F 33 are conspicuously good for transitron action.

Figure 16.6 shows the circuit. When it is quiescent, the suppressor potential is such as to encourage a substantial anode current and a much smaller

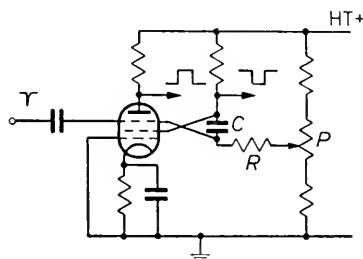


Figure 16.6

screen current, as in normal pentode usage. If now a negative trigger pulse be applied to the suppressor the anode current is reduced and electrons flow instead to the screen, increasing the screen current and reducing the screen potential. This reduction is transferred via C to the suppressor, producing a further fall in anode current and increase in screen current. Cumulative action then occurs as a result of which all the cathode current goes to the screen and the anode current is cut off. This state of affairs persists until, as a result of re-adjustment of the charge on C , the suppressor potential returns exponentially to a point at which anode current once more begins to flow. The screen current then falls, the screen potential rises and a reverse cumulative action occurs. The waveforms are similar to Figure 16.4, except that screen potential should replace V_1 anode potential, and suppressor potential the V_2 grid potential.

Uses for single square waves—The pulses produced by the flip-flop and transitron may be used for applications where their shape is specifically applicable, as in square wave stimulation, or use may primarily be made of the *duration* of the wave for the production of delays. If the square wave is passed through a CR network of short time constant*, then a pip is generated at the beginning and end of the wave, and the second pip selected by an appropriately connected diode, as in Figure 16.7. The second pip is then used to trigger further circuits.

Flip-flop circuits are so commonplace that there is little point in trying to refer the reader to all possible uses for them. The triggered transitron is

* Sometimes called a 'differentiating' circuit, since the output is responsive to rates of change at the input, rather than to the input itself.

TRIGGERED PULSE GENERATORS

less common. Attree¹ describes a stimulator employing a transitron delay stage and—incidentally—a flip-flop to form the stimulus pulses.

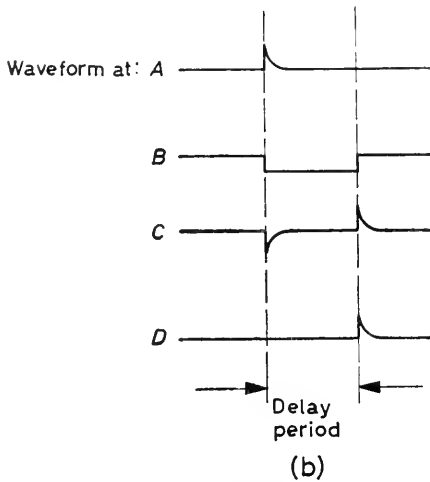
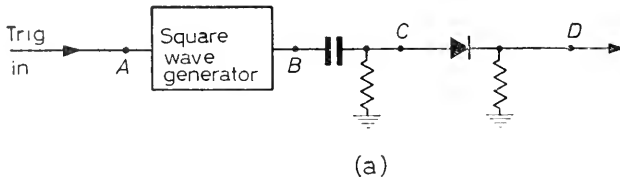


Figure 16.7

Eccles-Jordan circuit

This circuit is included for lack of anywhere better to put it. It is a square wave generator and it *does* require triggering, but whereas the flip-flop and transitron bring their square waves to an end automatically after a time determined by the charging of a capacitor, the Eccles-Jordan circuit requires a further trigger pulse to return it to the initial condition. It is, in fact, a bi-stable or two-state circuit, like the difference diode circuit in *Figure 7.9*.

In its simplest form it appears as in *Figure 16.8*. Two valves are connected to each other so that when both amplify there is much too much positive feedback for stability, and by cumulative action one valve is cut off and the other conducts hard. The resistance chains are chosen so that if V_1 is conducting hard, V_2 grid is held just negative enough to cut off V_2 and vice versa; the circuit is quite symmetrical, and will remain indefinitely in either state until the bias on the cut-off valve is lifted by a positive trigger pulse on its grid, when the circuit snaps over to the opposite condition.

The change-over speed may be improved by shunting the resistance from each anode to opposite grid by a capacitance, and by making the valves pentodes (*Figure 16.9*). The negative supply may be dispensed with by the use of automatic bias. This is not a cathode coupled circuit—either one valve

TRIANGULAR WAVE GENERATORS

or the other is conducting and if both consume the same current the potential difference across the cathode resistor is quite steady (*Figure 16.10*).

Eccles-Jordan circuits occur in designs of electrobiological interest by Dickinson² (time base), Attree¹ (stimulator) and many others.

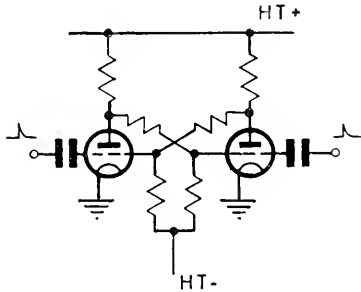


Figure 16.8

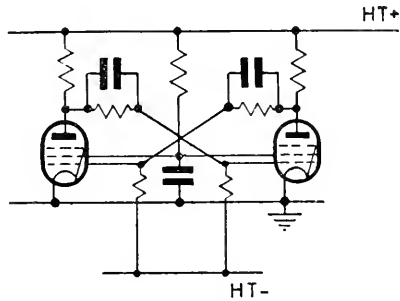


Figure 16.9

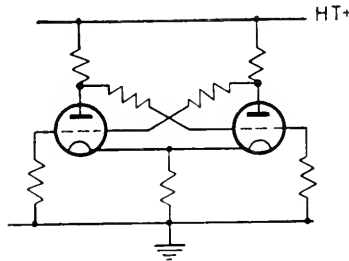


Figure 16.10

TRIANGULAR WAVE GENERATORS

The most important use for triangular waves is the production of time bases for cathode ray tube displays. Both electromagnetic and electrostatic methods of cathode ray deflection are relevant to electrobiology, so we have to consider generators of triangular waves of voltage, and triangular waves of current.

Triangular voltage waves

The basis of all triangular voltage wave production is the charging of a capacitance by a constant or nearly constant current, for then we have

$$V = \frac{Q}{C} = \frac{It}{C}$$

At the end of the charging period the capacitance is rapidly discharged, completing the waveform (*Figure 16.11*). Suitable sources of constant current are the saturated diode or the pentode (*Figure 16.12*). The latter method is adopted in the famous Puckle time-base generator³. The Puckle time base has certain disadvantages which render it unsuited to electrophysiological work, and it will not be discussed here.

TRIGGERED PULSE GENERATORS

If the capacitance is merely charged via a resistance, instead of a constant-current device, the waveform across it rises exponentially rather than linearly. We have already discussed three circuits—the difference diode relaxation oscillator, the thyatron relaxation oscillator and the blocking oscillator—

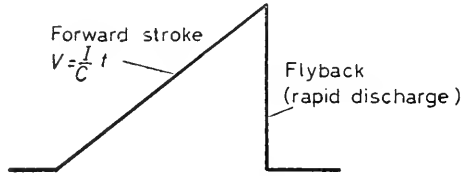


Figure 16.11

which produce waveforms of approximately triangular—in fact exponential—form. The degree to which the exponential produced resembles the ideal form depends on how much of it is used; if the capacitance is allowed to charge only to 1 per cent or so of the total charging voltage available

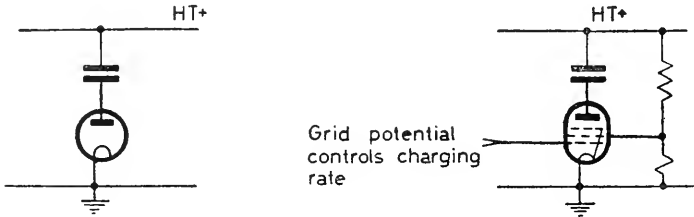


Figure 16.12

before discharge occurs, then the departure from linear rate of voltage rise is usually quite negligible (Figure 16.13). In order to produce a triangular wave—using resistance charging—of amplitude 200 V, we should need to be able to call upon a charging e.m.f. of 20,000 V; in general this is not

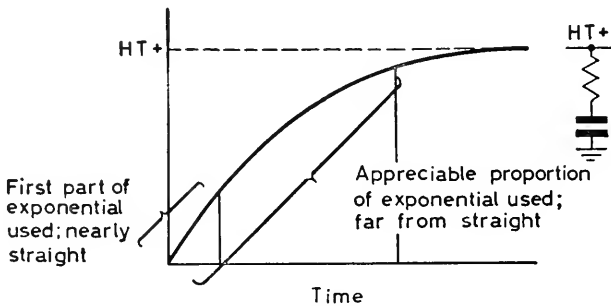


Figure 16.13

possible. However, we could produce a 2 V wave from a 200 V supply, and amplify it with a valve giving a gain of 100 times; but in practice there is a better way, which is to use a Blumlein integrator.

TRIANGULAR WAVE GENERATORS

Blumlein integrator

Consider the circuit in *Figure 16.14*, and suppose that the gain between

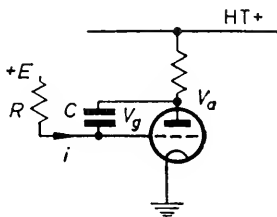


Figure 16.14

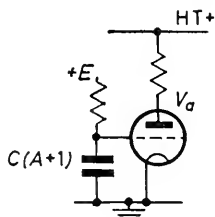


Figure 16.15

grid and anode is $-A$. Then we have, assuming no grid current and that therefore all the current i flows into C ,

$$i = -C \frac{\delta}{\delta t} (V_a - V_g)$$

also

$$\delta V_a = -A \delta V_g$$

$$i = C(A + 1) \frac{\delta V_g}{\delta t}$$

This equation describes what would happen if C were replaced by $C(A + 1)$ between grid and earth as in *Figure 16.15*, which is in effect a re-statement of the Miller effect, and the reason why the Blumlein integrator is often called a Miller integrator. We can now write down by inspection that the equation to V_g is

$$E(1 - e^{-\frac{t}{RC(A+1)}})$$

and therefore that

$$V_a = -AE(1 - e^{-\frac{t}{RC(A+1)}})$$

That is, it is as if a capacitance $(A + 1)$ times bigger than C were being charged by a voltage A times bigger than E . If A is 100 and E is made the full HT voltage, perhaps 350 V, then the virtual charging voltage is 35,000, which means that V_a will fall in a manner highly linear with time (called the Miller run-down). Moreover, since the virtual capacitance is 101 times the actual one, slow time bases can be generated with quite moderate values of C .

The 'Miller run-down' proceeds* in a linear manner at the anode until the anode voltage has fallen so low that A begins to fall. In order to secure a large amplitude of run-down it is customary to exploit the ability of a pentode to operate satisfactorily at low anode voltages, and if this is done the run-down is conveniently controlled at the suppressor grid. When the circuit is quiescent, the suppressor is biased sufficiently negative to cut off the anode current. To initiate the run-down the suppressor bias is lifted by

* Notice the run-down is preceded by an abrupt change of potential at grid and anode. This is necessary to carry the grid cathode potential from the grid current region to the amplifying region. 'Vertical fall' at the beginning of the stroke is characteristic of Miller time bases.

TRIGGERED PULSE GENERATORS

a positive-going square wave applied from elsewhere until an output of sufficient amplitude has been obtained (*Figure 16.16*). On cutting off anode current once more, the anode tends to rise in potential taking the control grid with it—via C —and causing grid current to flow. This current recharges C and allows the anode to return to HT plus potential with time constant $R_L C$. R is much greater than R_L and plays no significant part in the flyback operation.

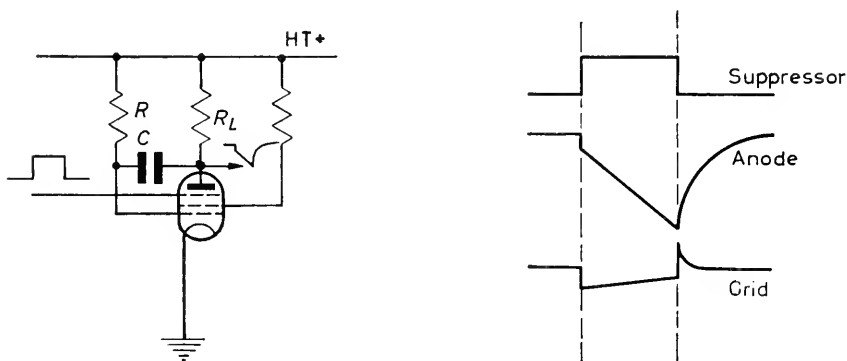


Figure 16.16

Whilst it is clear that the forward stroke time constant $RC(A + 1)$ is enormously greater than the flyback time constant $R_L C$, we usually use so little of the former that the flyback time is often of the same order as the forward stroke time (the two times are roughly in the ratio R_L/R) which is undesirable, since one wants the circuit to be available for a further forward stroke, if required, as quickly as possible. We shall mention now two refinements to the Blumlein integrator, one of which aims at improving matters here. It is due to Attree⁴.

*Attree refinement*⁴—Here a cathode follower is interposed between the pentode anode and the upper plate of C . This has almost no effect on the operation of the circuit during the forward stroke. On the flyback the anode is at liberty to return quickly to HT plus, the charge on C being replaced quickly through the low output resistance of the cathode follower, perhaps 500 ohms, instead of via R_L , perhaps 50,000 ohms. The flyback is thus 100 times quicker (*Figure 16.17*).

*Gibbs and Rushton refinement*⁵—When it is required to produce a triangular wave lasting for periods of the order of 5 minutes it is undesirable to achieve such protracted forward strokes by using enormous values of R and C , since large capacitors are bulky and expensive, and high resistances of upwards of 1 k megohms are also expensive and possibly of dubious stability, becoming comparable with the circuit leakage resistances. Let us illustrate with an example. Differentiating the expression for V_a ,

we get

$$\frac{dV_a}{dt} = -\frac{AEt}{RC(A + 1)} e^{-\frac{t}{RC(A + 1)}}$$

TRIANGULAR WAVE GENERATORS

and because $t \ll RC(A + 1)$ in the working range, and $A \gg 1$

$$\frac{dV_a}{dt} \approx \frac{tE}{RC}$$

So the time to generate a given output amplitude V_{out} is $t = CR V_{out}/E$.

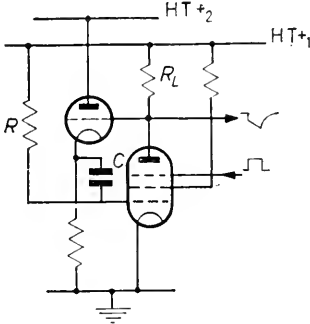


Figure 16.17

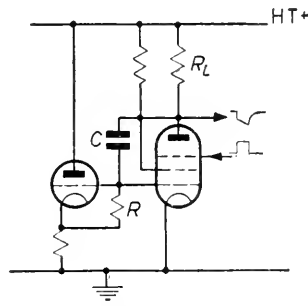


Figure 16.18

If E is derived from the HT supply and is, say, 250 V, and V_{out} is, say, 200 V, and $t = 300$ seconds

$$CR = \frac{300 \times 250}{200} = 375 \text{ megohm-microfarads}$$

The solution is of course to reduce E , but if we do this we reduce the virtual charging voltage AE and spoil the linearity. The Gibbs and Rushton refinement makes use of a 'bootstrap' circuit (Figure 16.18). Any cathode-follower cathode will set itself up a few volts positive to its grid. Applied to this circuit, these few volts constitute our charging voltage, E . Current flows into C via R from V_1 cathode, but the concomitant small rise in grid potential as the stroke proceeds does not reduce the rate of charging because the grid potential rise is transferred to V_1 grid and produces a similar increase in E , maintaining the charging current. If we can by this method, without impairing the linearity, reduce E from 250 V to 2.5, we can secure a 5 minute sweep with $CR = 3.75$ megohm-microfarads, say $C = 0.1$ microfarad and $R = 37.5$ megohms, both reasonable values.



Figure 16.19

Control of Miller run-down—We have now to discuss how to derive the square wave which allows the Miller run-down to occur. There are a number of possible ways, depending on what is wanted.

If the requirement is for independent control of the duration and slope of the wave (from which it follows that the amplitude will vary) then all that is necessary is to precede the Blumlein integrator with a flip-flop. The flip-flop

TRIGGERED PULSE GENERATORS

time constant will then control the duration, and the integrator time constant the slope (*Figure 16.19*). More often the requirement is to vary the duration whilst maintaining the amplitude. This can be done with the above arrangement operating the two controls together, but this is inconvenient and can be avoided by the use of one of the special circuits which follow.

Eccles-Jordan control

This circuit (*Figure 16.20*) appears to be due to Dickinson². Normally V_1 is conducting and V_2 cut off. V_1 anode potential is therefore low and the values of the potential divider resistances R_1 and R_2 , in conjunction with the negative supply, are such that V_3 anode current is just cut off by the suppressor.

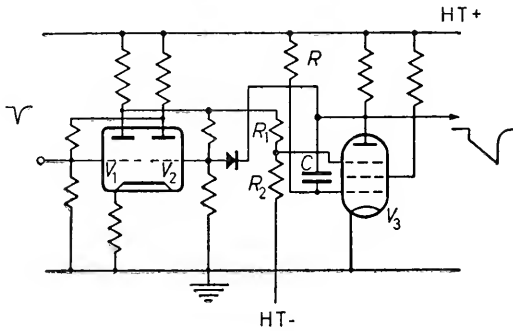


Figure 16.20

Upon the arrival of a negative trigger pulse the Eccles-Jordan circuit 'turns over' and V_1 anode goes positive, lifting the suppressor bias on V_3 and allowing the Miller run-down to begin. This proceeds until V_3 anode reaches a point at which the diode conducts, making V_2 grid negative and returning the Eccles-Jordan circuit to the original state. Anode current in V_3 is then cut off once more by the suppressor and the flyback occurs. The amplitude of the output is clearly independent of the rate of Miller run-down.

Sanatron

In this circuit⁶ (*Figure 16.21*) the Miller valve is controlled by a single pentode. V_1 is normally conducting hard and V_1 anode potential is low; R_1 and R_2 and the negative supply are arranged so that anode current in V_2 is just cut off by V_2 suppressor.

The trigger pulse cuts off V_1 on its suppressor, allowing V_1 anode to go positive and removing the bias on V_2 suppressor. Miller run-down begins, generating via C' an approximately constant negative potential across R'^* , so that V_1 remains cut off by its control grid. This state of affairs persists until the run-down begins to slow, and the negative bias, holding V_1 cut off, begins to fall. V_1 then conducts once more, switching off anode current in V_2 and initiating the flyback.

The circuit will give a constant amplitude of output provided $C'R'$ and

* Because $iR' = C'(d/dt)(V_a)$, which is constant during the run-down.

TRIANGULAR WAVE GENERATORS

CR are altered together. If $C'R'$ is too small the circuit will not work at all, or will give only a small output. If $C'R'$ is too big the output has the form of *Figure 16.22*.

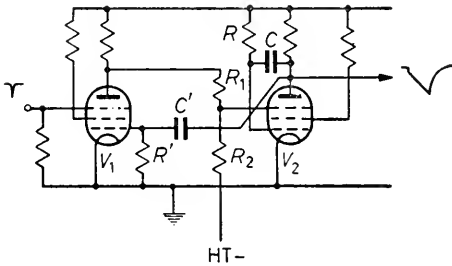


Figure 16.21

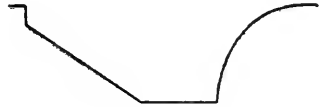


Figure 16.22

Phantastron

In this circuit one valve is used to its utmost; not only all three grids and the anode, but also the cathode fluctuate in potential to contribute to the action. Stripped of unessentials, the circuit has the appearance of *Figure 16.23*. R_1 , R_2 and R_3 are merely a potential divider across the HT supply to

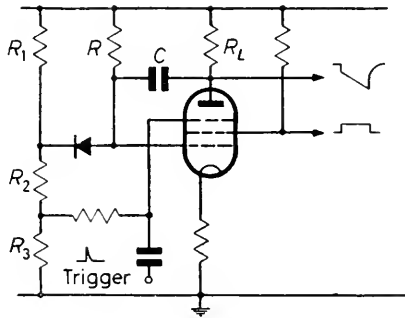


Figure 16.23

define the potentials of suppressor and control grid when the circuit is quiescent. Notice the catching diode, which allows the control grid to move negative with respect to the junction of R_1 and R_2 , but not positive.

Before the arrival of the trigger pulse screen current flows and by cathode follower action the cathode is slightly positive with respect to the control grid. As a result of the voltage drop along R_2 the suppressor is considerably more negative and in consequence anode current is cut off. On the arrival of a positive trigger, the suppressor potential is raised, allowing anode current to flow. Vertical fall occurs at the anode and is communicated to the control grid via C and to the cathode by cathode-follower action. Since control grid and cathode have fallen it is as if the suppressor potential had been raised, and Miller run-down can proceed even though the trigger pulse has come to an end.

TRIGGERED PULSE GENERATORS

Miller run-down continues, the anode current rising steadily, until the potential difference developed across the cathode resistor by the rising anode current once more makes the cathode sufficiently positive to cut off anode current at the suppressor. Flyback then occurs. Phantastron waveforms

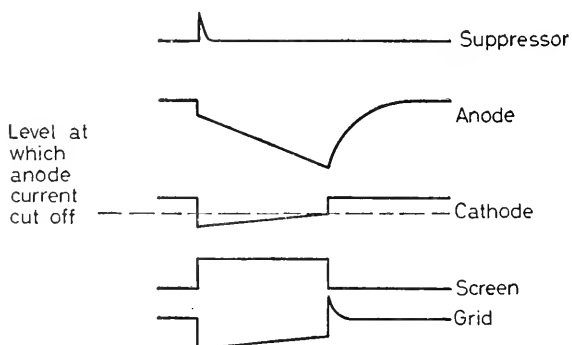


Figure 16.24

have the form of *Figure 16.24*. The Phantastron may also be triggered by negative pulses applied to anode or control grid.

Triggered Miller transitron

This extremely simple circuit (*Figure 16.25*) is a combination of the Blumlein integrator and transitron, but it should be noted that the sense of the transitron bias is the reverse of that used in the transitron square wave generator.

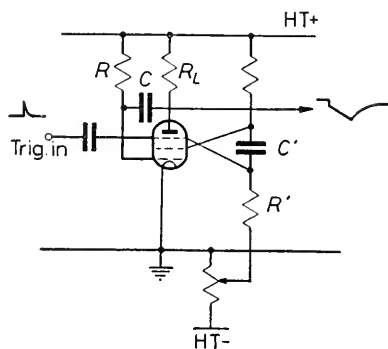


Figure 16.25

In the latter the suppressor bias is, if anything, rather positive, so that in the quiescent state the anode current is large and the screen current small. In this circuit the suppressor is biased so that anode current is just cut off.

A positive trigger pulse applied to the suppressor allows anode current to flow, reducing the screen current. There is thus a rise in screen potential which is passed on to the suppressor via C' , so that the suppressor bias continues to be held off after the trigger pulse ends. Miller run-down ensues,

TRIANGULAR WAVE GENERATORS

accompanied by continued rise in screen potential and holding off of suppressor bias until the run-down can proceed no further. Suppressor bias then re-asserts itself and the anode current is cut off; flyback occurs.

A time base using a triggered Miller transistor has the advantages of simplicity and the ease with which it can be made a 'free-running' generator of continuous triangular voltage oscillations, merely by reducing the suppressor bias. Its snag lies in its rather long refractory period between strokes; this is bound up with the ratio of $R'C'$ to RC . If $R'C'$ is large enough to ensure the continued holding off of suppressor bias, then after flyback, when the screen moves negative once more, there is a rather long period of extreme suppressor negativity which the trigger pulse will be unable to overcome. This is clear from the waveforms in *Figure 16.26*.

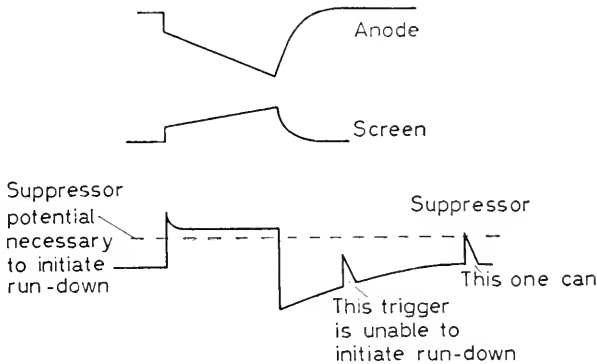


Figure 16.26

Trace brightening—It is worth noting that all the triggered triangular wave generators listed generate somewhere a positive going square wave. Where these circuits are used for cathode ray time bases this wave may usefully be applied either through a CR circuit or through an RF coupling device (see Chapter 32) to brighten the beam during its traverse, the beam being suppressed at other times.

Triangular current wave generators

In this section we are concerned with the production of triangular waves of current in an inductive device such as a beam-deflector coil for a cathode ray tube. Such a coil comprises inductance and resistance in series (*Figure 16.27*), the resistance being of course the inevitable resistance of the wire with which the coil is wound.

If it is required to produce a current which rises at a steady $di/dt = \dot{I}$, then the potential difference across the coil terminals required to produce it has the 'penthouse' form of *Figure 16.28* and comprises two components: a square wave to overcome the back e.m.f., of magnitude $L\dot{I}$, plus a triangular wave of slope $\dot{I}R$, to overcome the drop across the resistance.

There are thus two approaches to the problem of triangular current generation: to supply a current wave of the form of *Figure 16.28a* from a constant-current generator, or to supply a voltage wave of the form of

TRIGGERED PULSE GENERATORS

Figure 16.28b from a constant-voltage generator. We take as an example a deflector coil requiring 50 mA to sweep the cathode ray beam across the full diameter of the screen (assuming that by some external and unspecified means the spot is caused to rest at one edge of the screen when the

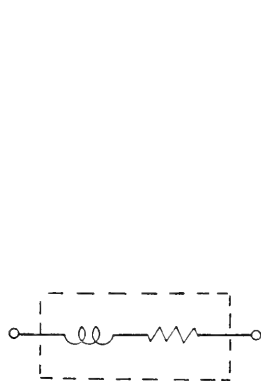


Figure 16.27

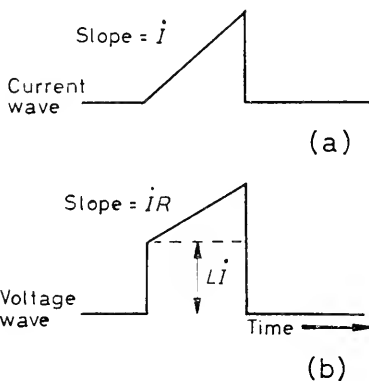


Figure 16.28

deflector coil current is zero) and traverses to the diametral edge when the current is 50 mA. Let the inductance be 3 H and the resistance 100 ohms, and assume that the fastest time base required is 1 millisecond and the slowest 1 second.

Constant-current approach

The relevant constant-current generator here is clearly an output pentode, with the deflector coil in the anode circuit and a triangular voltage wave

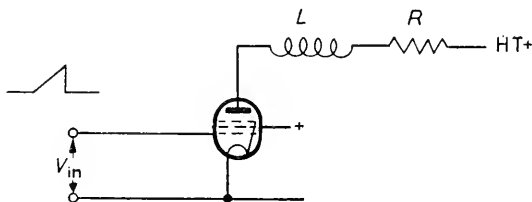


Figure 16.29

applied to the grid (Figure 16.29). To secure economy in anode current the input wave will be positive going, so that the anode current in the quiescent state is small and determined by the grid bias obtained from the datum potential of the input wave. The working point on the anode characteristic will be *A* in Figure 16.30. To get an anode current of 50 mA we need to swing the grid positive an amount $50 \text{ mA}/g_m$.

Suppose we are using the slowest sweep speed. \dot{i} is 50 mA in 1 second, or 0.05 amps/second and the back e.m.f. will be $3 \times 0.05 = 0.15 \text{ V}$, quite negligible. We can thus forget the effect of the deflector coil inductance and say that during the forward stroke of the circuit the working point will move along a rather steep load line (corresponding to 1,000 ohms—rather small as loads go) from *A* to *B*.

TRIANGULAR WAVE GENERATORS

On the fastest sweep speed matters are rather different. \dot{I} is 50 milliamps in 1 millisecond = 50 amps/sec and the back e.m.f. is therefore $3 \times 50 = 150$ V. This back e.m.f. becomes active as soon as the forward stroke begins, sweeping the working point 150 V to the left to A' . Thereafter it moves to

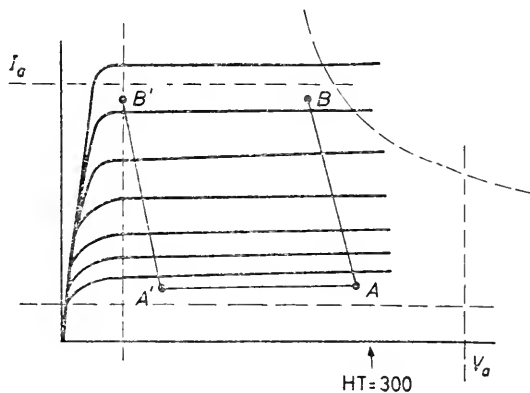


Figure 16.30

B' along a path parallel to AB , as the voltage drop across the resistance increases. At the end of the sweep the total deflector coil voltage is $L\dot{I} + \dot{I}_{\max}R = 200$ V, leaving an anode voltage for the valve of 100 V.

From the foregoing the points to watch in design emerge; they are:

- (1) B must not move outside the maximum power hyperbola;
- (2) B' must not move outside the working region at the knees of the anode characteristics;

(3) neither B nor B' must go above the maximum anode current line.

The parallelogram $ABB'A'$ is a property of the particular deflector coil used and the upper and lower sweep speeds. It may be drawn on tracing paper and juggled about on the anode characteristics for various valves until one is found which it will neatly fit. The quiescent bias and HT+ voltage required are then determined. If the parallelogram cannot be made to fit because P_{\max} or $I_{a-\max}$ for the valve is infringed, consider using two valves in parallel. This halves the current each has to supply and also, incidentally, since g_m has been doubled, halves the amplitude of input wave required at the grid. If it cannot be made to fit because no valve will stand the amount of HT required, a new deflector coil having fewer turns is indicated. For example, if we halve the number of turns, the inductance L will be reduced to one quarter, \dot{I} will have to be doubled, so the back e.m.f., $L\dot{I}$, will be halved. The input wave amplitude will also in general require doubling. The extra current required by the new coil may require two valves in parallel in place of one, in which case the input required remains the same.

Negative current feedback

The linearity of the output current depends on the constant-current generator living up to its name. The pentode is entitled to be regarded as a constant-current generator in virtue of its high internal resistance r_a , of the order of a

TRIGGERED PULSE GENERATORS

megohm. The internal resistance can be made even higher by the use of negative *current* feedback. In physical terms this means feeding back a voltage proportional to the output current, rather than to the output voltage. It will be remembered that the output resistance of the cathode follower, which represents negative voltage feedback par excellence, is very low. With negative current feedback the output impedance is very high. A simple

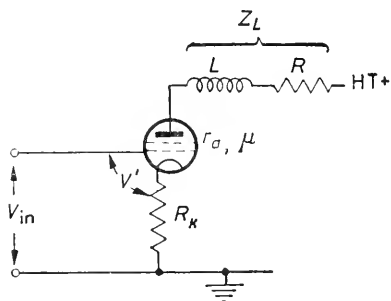


Figure 16.31

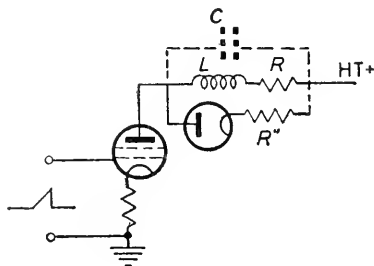


Figure 16.32

way of achieving negative current feedback is to include some resistance in series with the valve cathode (Figure 16.31). Then we have

$$i = \frac{V' \mu}{Z_L + R_K + r_a} \quad \text{and} \quad V' = V_{in} - iR_K$$

Combining

$$i = \frac{\mu V_{in}}{Z_L + r_a + (\mu + 1)R_K}$$

Comparing this with the expression for a stage without feedback having the same load and a valve of internal resistance r'_a

$$i = \frac{\mu V_{in}}{Z_L + r'}$$

we see that the effect of the feedback is to increase the internal resistance from r_a to $r_a + (\mu + 1)R_K$. This is seen better if we call it

$$r_a \left(1 + \frac{\mu + 1}{r_a} R_K \right) \cong r_a (1 + g_m R_K)$$

Typically g_m is 10 mA/V and r_k 1,000 ohms, giving an 11-fold increase in the internal resistance of our constant-current generator. It must not be forgotten that the improvement is not obtainable without a loss of gain. The input has to be increased by a factor

$$\frac{Z_L + r_a + (\mu + 1)R_K}{Z_L + r_a + R_K} \approx \frac{r_a + (\mu + 1)R_K}{r_a} \approx g_m R_K$$

i.e. approximately by the same factor as the pentode internal resistance is increased.

Flyback—During the forward stroke energy is stored in the magnetic field linked with the deflector coil, and when the anode current is suddenly

TRIANGULAR WAVE GENERATORS

reduced at the flyback this field collapses, causing the back e.m.f. to reverse and tend to maintain the anode current by driving the pentode anode very positive. In view of the high rate of change of anode current this e.m.f. is very large and may damage the valve or the coil insulation unless limited.

The procedure is to use a diode as a switch to connect a resistance automatically across the coil during the flyback, so that the magnetic energy is dissipated as quickly as possible in the resistance as heat. If R is neglected, the coil and its stray capacitance form a simple parallel-tuned circuit, and the resistance will introduce shunt damping. Reference to Chapter 5 shows us that the energy will be removed quickest when the circuit is critically damped, i.e. when $R' = \frac{1}{2}(L/C)^{\frac{1}{2}}$. The existence of R means that some series damping is already present, and the critical value of R' will be somewhat higher. The scheme is shown in *Figure 16.32*.

Source of triangular wave—Any of the triangular voltage wave generators may be used to feed the pentode, but as their outputs are mostly negative-going, a buffer amplifying stage to provide phase-inversion is generally required.

Constant-voltage approach

In this case the deflector coil may be fed from a cathode follower as in *Figure 16.33*, the appropriate penthouse wave being applied to the grid. It is in generating this appropriate wave that a complication with this method arises, since a different amplitude of square wave component is needed for

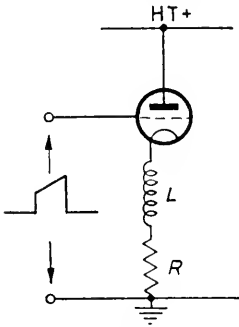


Figure 16.33

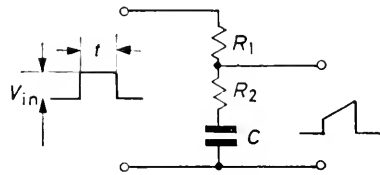


Figure 16.34

each sweep speed. Mixing resistors may be used to combine the outputs of a flip-flop, and a Blumlein integrator controlled by it, in correct proportions, and it is wise not to aim at continuously variable speed control but to be content with switched speeds in discrete steps. The same control can then switch in different mixing resistors as required. Alternatively, approximate penthouse waveforms can be generated from a square wave by a circuit of the type of *Figure 16.34*. If the time constant $C(R_1 + R_2)$ is much longer than t , the penthouse rise is nearly linear, and then

$$\text{square wave component} = V_{\text{in}} \frac{R_2}{R_1 + R_2}$$

$$\text{triangular wave slope} = \frac{V_{\text{in}}}{C(R_1 + R_2)}$$

TRIGGERED PULSE GENERATORS

In view of the complications involved in correct penthouse waveform generation the author has never used constant-voltage deflector coil drive. Its advantage lies in its management of a quantity we have rather ignored, the coil stray capacitance. The cathode follower has no difficulty in supplying extra current to charge this, whilst still maintaining the correct P.D. across the coil (*Figure 16.35*). The pentode, on the other hand, is obliged to supply a

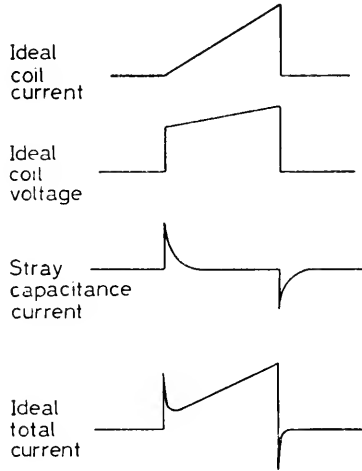


Figure 16.35

triangular *total* current, which means that the coil current cannot itself be triangular, there must be some distortion. In the author's experience, with coils made with reasonable care the effect is not serious.

REFERENCES

- ¹ ATTREE, V. H. *J. sci. Instrum.* 27 (1950) 43
- ² DICKINSON, C. J. Electronic engineering monograph *Electrophysiological Technique* p. 49
- ³ PUCKLE, O. S. *Time Bases* London; Chapman & Hall
- ⁴ ATTREE, V. H. *J. sci. Instrum.* 26 (1949) 257
- ⁵ GIBBS, D. F. and RUSHTON W. A. H. *J. sci. Instrum.* 23 (1946) 220
- ⁶ MOODY, N. F. and WILLIAMS F. C. *J. Instn. elect. Engrs.* 93 Pt IIIA, No 7

NOISE

It is not possible to record an indefinitely small signal by the employment of indefinitely large amounts of amplification. This is because there are generated within the source of the signal itself, and within the amplifying apparatus, random electrical effects of a definite statistical magnitude. If the effect produced at the indicating device, meter, cathode ray tube, etc., by this random activity is of such a size as to obscure the effect produced by the signal, the latter will at least require special measurement techniques, and may be impossible to measure at all.

This electrical disturbance is called 'noise', a most apt description for it when heard over a loudspeaker: the sound produced may be described as 'rushing'. Because of its random nature nothing can be said about its instantaneous amplitude or frequency. It is, however, possible to allot it a root mean square amplitude and a frequency spectrum, which is the frequency spectrum or band-width of the recording gear.

The classical causes of noise are 'resistance noise' and 'valve noise'. The former is due to the mechanical oscillations of the molecules of which the resistance is composed; the latter is due to the inhomogeneous nature of the electron stream. When one tries to frame a definition in general terms, such as 'noise is produced wherever there is molecular agitation, or where a flow of current is of an irregular nature', a great many other deficiencies of recording gear become difficult to exclude—for example, who can say that the current from a dying HT battery or the supply mains at winter rush hour is not irregular? In this way the concept of 'noise' has tended to broaden, so that in some quarters 'noise' is held to be 'anything which is not signal'. We shall confine the term noise to its narrow sense, and where other spurious effects have a name we shall use it.

HOMOGENEOUS RESISTANCE NOISE (OTHERWISE THERMAL OR JOHNSON NOISE)

The R.M.S. noise power delivered into a matched load by any homogeneous piece of material between two terminals is given by

$$kTB$$

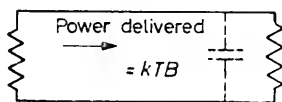


Figure 17.1

where k is Boltzmann's constant, T is the absolute temperature and B is the band-width of the device in c/s—the upper frequency in this case being limited by the shunting effect of stray capacitance (Figure 17.1). It is worth

NOISE

noticing that the nature of the material itself does not enter into the calculation. Notice also that we do not have here any kind of perpetual motion device; the load supplies an equal power back to the source and there is no net transfer of energy.

If R is the resistance of the noise source and V_L is the R.M.S. voltage delivered to the load, we have

$$\text{Noise power} = \frac{V_L^2}{R}$$

$$V_L^2 = kTBR$$

$$V_L = (kTBR)^{1/2}$$

and the R.M.S. *open circuit* noise voltage will be twice this

$$V_{oc} = 2(kTBR)^{1/2}$$

$k = 1.37 \times 10^{-23}$ joules per degree absolute, so the open circuit R.M.S. noise voltage generated by a 1 megohm wire-wound resistance at 24°C as measured via an amplifier of band-width 0–10 kc/s is

$$\begin{aligned} & 2(1.37 \times 10^{-23} \times 3 \times 10^2 \times 10^4 \times 10^6)^{1/2} \\ &= 2(411 \times 10^{-11})^{1/2} \\ &= 12.8 \text{ microvolts} \end{aligned}$$

CARBON RESISTOR NOISE

The most commonly used resistors in electronics are not composed of a homogeneous material; they are an aggregate of carbon particles bound together, and through molecular agitation these particles continually vary in the intimacy of their contact one with another, as a result of which the overall resistance of the component is itself subject to continual minor fluctuation. Thus a steady current through the resistor causes a fluctuating potential difference to appear across it. Noise caused in this manner is over and above ordinary Johnson noise, and is dealt with again in Chapter 20. For the present it is sufficient to bear in mind that carbon resistor noise is a function of the current the component is carrying.

VALVE NOISE

Shot effect

This is caused by the random manner in which electrons are emitted from the cathode. It is most pronounced in the saturated thermionic diode, so much so that saturated diodes are commonly used in certain pieces of test apparatus whose function it is to generate noise. It can be shown¹ that the R.M.S. noise voltage appearing at the anode of such a diode is given by

$$V = (2IeB)^{1/2} \times R$$

where I is the diode current, e is the charge on an electron $= 1.6 \times 10^{-19}$ coulombs, B is the band-width in c/s and R is the diode load. Thus if the diode anode current $= 1$ mA, the band-width is 10 kc/s and the anode load is 100 k Ω , the R.M.S. noise voltage produced is

$$\begin{aligned} V &= (2 \times 10^{-3} \times 1.6 \times 10^{-19} \times 10^4)^{1/2} \times 10^5 \\ &= 179 \text{ microvolts} \end{aligned}$$

VALVE NOISE

If the diode is operated under space-charge limiting conditions the noise generated falls. This may be explained by a 'smoothing out' of anode current by the space charge. Under these conditions the noise voltage is given by the Johnson noise that would be caused by a resistance, equal to the diode incremental resistance r_a , raised to a temperature 0.64 of that of the diode cathode, i.e.

$$V_{oc} = 2(0.64 T_{cath} kB r_a)^{1/2}$$

and the proportion of this which appears across an anode load R is

$$V_a = \frac{R}{R + r_a} \cdot 2(0.64 T_{cath} kB r_a)^{1/2}$$

Thus taking $R = 100 \text{ k}\Omega$, $r_a = 50 \text{ k}\Omega$, $T_{cath} = 1,000$ absolute, $B = 10,000 \text{ c/s}$ we get

$$\begin{aligned} V_a &= \frac{100}{100 + 50} \times 2 \cdot (0.64 \times 10^3 \times 1.37 \times 10^{-23} \times 10^4 \times 5 \times 10^4)^{1/2} \\ &= 2.8 \text{ microvolts} \end{aligned}$$

If now a grid be introduced and the valve become a triode, the generated noise is very slightly raised. It can be shown that the expression for space-charge limited diode noise voltage just given should be multiplied by a factor $\{1 + (1/\mu)\}^{1/2}$.

Partition noise

With tetrodes and pentodes additional noise is caused by the random fashion in which the cathode current divides between the screen and the anode. This is called the 'partition noise'. The space charge does not extend out to the partition region, so the expression for partition noise is related to the expression for a diode under saturated conditions. It follows from equations derived by Starr that the R.M.S. partition noise voltage appearing at the anode of a pentode or beam valve is

$$V_p = \left(2 eB \frac{I_{sg} \cdot I_a}{I_a + I_{sg}} \right)^{1/2} R$$

where I_{sg} is the screen grid current and I_a is the anode current, R is the anode load and e and B have the usual meanings. Thus if $I_a = 0.8 \text{ mA}$, $I_{sg} = 0.2 \text{ mA}$, $B = 10 \text{ kc/s}$, $R = 100 \text{ k}\Omega$, the R.M.S. partition noise is

$$\begin{aligned} V_p &= \left(2 \times 1.6 \times 10^{-19} \times 10^4 \times 0.8 \times 10^3 \times \frac{0.8 \times 0.2}{0.8 + 0.2} \right)^{1/2} \times 10^5 \\ &= 6.8 \text{ microvolts} \end{aligned}$$

Flicker effect

This is a very slow noise-like phenomenon of peculiar interest to electro-physiologists, whose amplifier pass-bands commonly extend to lower frequencies than those of most electronic engineers. Flicker noise, unlike Johnson, shot and partition noise, is not distributed evenly throughout all

frequencies; the flicker noise voltage in a small band $\delta\omega$ has been shown by Macfarlane² to be proportional to $(1/\omega)^x$ where x lies between $\frac{1}{2}$ and 1. It is also proportional to $(I_k)^y$ where y is between 1 and 1.5, I_k being the cathode current of the valve.

Microphony

Microphony is a modulation of valve anode current caused by small variations in the geometry of the valve structure. These variations may be produced by the valve itself by 'settling' of its structure as it warms up, or may be caused by vibrations from an external source—usually the experimenter—being communicated to the valve via its mounting arrangements or through the air. Microphony is combated by keeping the anode current low, choosing valves specifically made for non-microphonicity (by giving them a specially braced structure), keeping still and, when all else fails, resilient mounting and enclosure in a wadded container.

Hum and Drift—These will be dealt with in Part IV.

ORDERS OF MAGNITUDE

It is not hard to see that as the noise generated in the low-level part of an amplifier receives the most subsequent amplification, efforts to reduce noise must primarily be directed at the first amplifying stage and, where there is any freedom of choice, the signal source itself. It is clear from the expressions that have been given that valve noise may be minimized by low heater voltage (shot noise) and low anode and screen voltages (partition noise, flicker, microphony) and by correct choice of valve (triode if possible, otherwise low ratio of screen to anode current, and of anti-microphonic construction).

Whether or not valve noise is important depends on the signal source. If the latter is an E.E.G. subject of internal resistance 10,000 ohms, and the band-width of the apparatus is 100 cycles, as it well might be, then at 24°C the R.M.S. noise voltage produced by the patient is 0.128 microvolt, and if the first stage gain is 30 the effect produced at the anode is 30 times greater, i.e. 3.84 microvolts. Correcting the valve noise in the examples given for the reduced band-width, we get a triode noise of about 0.3 microvolt and a pentode noise of about 0.7 microvolt, plus the effects of flicker. Valve noise is comparable in magnitude to the Johnson noise and worthy of attention.

On the other hand, suppose one is recording from a 10 megohm glass micro electrode over a band of 10 kc/s. Then the source noise cannot be less than 40 microvolts at 24°C at the first amplifying stage grid, or 1,200 microvolts at the anode. This is clearly enormously greater than valve noise over the same band, leaving one more free to choose the valve with other points in mind.

NEGATIVE FEEDBACK AND NOISE

The effect of negative feedback applied to an amplifier is to reduce the noise from all causes. Thus in *Figure 17.2* we have an amplifier of gain A , the noise content of whose output is supposed to be generated by the external generator NG ; n is the instantaneous noise voltage. Suppose now a fraction B be

CASCODE CIRCUIT

negatively fed back (*Figure 17.3*). Let the output of the device be instantaneously 1. Then the feedback amplifier input is B and the amplifier output is $-AB$. The instantaneous generator potential must therefore be $1 + AB$.

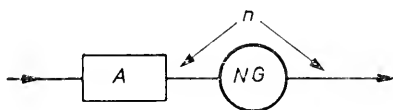


Figure 17.2

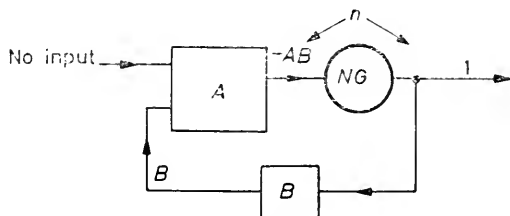


Figure 17.3

But this is by definition n , so the output in terms of n is $\frac{n}{1 + AB} \approx \frac{n}{AB}$, that is, the effect of the feedback is to make the noise in the output AB times less.

Cathode follower noise

An important application of the above is to the cathode followers which frequently precede the first amplifying valve in electrophysiology. Here B equals 1 and A is the gain, without feedback, $g_m R$, between 30 and 100. Thus the noise generated by the cathode follower is only between 1/30 and 1/100 of the noise it would produce if it were an amplifying stage. If the cathode follower is preceded by a glass microelectrode, as is usually the case, its noise is quite unimportant.

CASCODE CIRCUIT

There is an interesting circuit which has triode-like noise but pentode-like gain. It is called the cascode and is shown in *Figure 17.4*. Two triodes are used, usually similar, the grid of the upper one being returned to a fixed positive potential in a similar manner to a pentode screen grid. Both valves run under space charge limited conditions and so there is reduced shot noise from each but no partition noise. We have for the lower valve

$$\delta i = \frac{1}{r_a} \left\{ \delta V_{a_1} + \mu \delta V_{i_n} \right\} \quad \dots (1)$$

and for the upper valve

$$\delta i = \frac{1}{r_a} \left\{ \delta V_{a_2} + \mu \delta V_{g_2 k_2} \right\} = \frac{1}{r_a} \left\{ -\delta i R - \mu \delta V_{a_1} \right\} \quad \dots (2)$$

From 2

$$\delta V_{a_1} = \frac{-\delta i (R + r_a)}{\mu}$$

Substituting in 1

$$\delta i = \frac{1}{r_a} \left\{ \mu \delta V_{in} - \frac{\delta i(R + r_a)}{\mu} \right\}$$

$$\therefore \delta i \left(1 + \frac{R + r_a}{r_a \mu} \right) = \frac{\mu}{r_a} \delta V_{in}$$

$$\therefore \delta i = \frac{g_m \delta V_{in}}{1 + \frac{R + r_a}{r_a \mu}}$$

$$\therefore \text{The gain} = \frac{\delta i R}{\delta V_{in}} = \frac{g_m R}{1 + \frac{R + r_a}{r_a \mu}}$$

In a typical case $R = 30 \text{ k}\Omega$, $r_a = 10 \text{ k}\Omega$, $g_m = 5 \text{ mA/V}$, $\mu = 50$. Then $(R + r_a)/(r_a \mu) = 0.08$, which is $\ll 1$. Therefore gain $\approx g_m R$, as for a pentode.

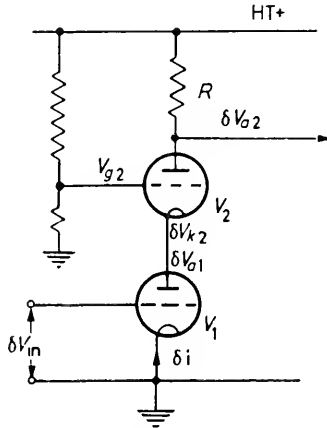


Figure 17.4

The cascode circuit also resembles the pentode circuit in that the input capacitance is low. The change in anode potential at the lower valve for a change of anode current δi is not $-\delta i R$ but $-\delta i \{(R + r_a)/\mu\}$. The Miller input capacitance is therefore reduced by a factor $\{(R + r_a)/\mu R\} = 1/37\frac{1}{2}$ of what it would be for a simple triode amplifier.

REFERENCES

¹ STARR, A. T. *Radio and Radar Technique* London; Pitman
² MACFARLANE, G. G. *Proc. phys. Soc. Lond.* 59 (1941) 366

THE CATHODALLY SCREENED CATHODE FOLLOWER

In electrophysiological work with glass micro-pipettes it is usual to make use of the high input impedance of the cathode follower and employ such a stage as a buffer between micro-electrode and amplifier. The piece of wire joining the electrode to the cathode-follower grid is extremely prone to pick up interference and requires screening.

In order that the capacitance between grid lead and screen shall have the minimum effect, the screen is connected to the valve cathode instead of to earth; if this capacitance is C_s and the cathode-follower gain is G , the effective input capacitance is not $C_s + (1 - G)C_{gk} + C_{ga}$, as it would be with the screen earthed, but only $(1 - G)(C_s + C_{gk}) + C_{ga}$. The technique is called cathodal screening and is due to M. Ryle.

To prevent interference pick-up by the cathodal screening, it is itself usually surrounded by an earthed screen (Figure 18.1). The presence of the

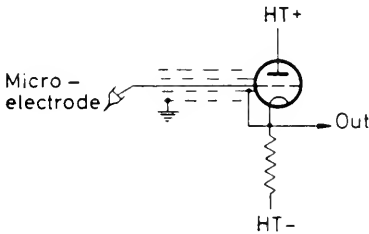


Figure 18.1

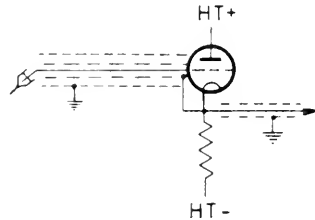


Figure 18.2

earthed screen has little effect on the performance, since it is in parallel with the low output-impedance of the cathode follower.

However, when use is made of the ability of the cathode follower to provide cathodal screening *and* to drive a long piece of cable, as in the typical electrophysiological circuit of Figure 18.2, it is possible to run into difficulties; it has been pointed out by W. J. Nastuk and A. L. Hodgkin (*J. cell. comp. Physiol.* 35 (1950) 39) that the transient response of the cathode follower can be oscillatory.

Steady-state response

The circuit may be drawn as in Figure 18.3. R_{in} represents the micro-electrode and C_s is the capacitance of the cathodal screen. C_o is the capacitance between the inner conductor of the output cable and its screen. The valve capacitances are likely to be relatively insignificant.

THE CATHODALLY SCREENED CATHODE FOLLOWER

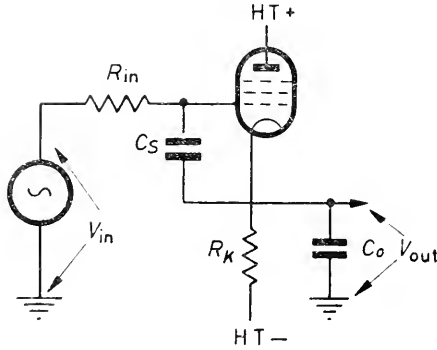


Figure 18.3

If the vector grid-cathode voltage is V , the valve current must be $g_m V$. The output voltage produced by this current is

$$g_m V \left(\frac{R_K \cdot 1/(j\omega C_o)}{R_K + 1/(j\omega C_o)} \right) \\ = g_m V \frac{R_K}{1 + j\omega R_K C_o}$$

The current through C_s must be $j\omega V \cdot C_s$, and if the grid current is negligible, this current must flow through R_{in} , producing across it a potential difference $j\omega V \cdot C_s \cdot R_{in}$. The total input voltage must therefore be

$$j\omega V \cdot C_s \cdot R_{in} + V + g_m V \frac{R_K}{1 + j\omega R_K C_o}$$

therefore

$$\frac{V_{out}}{V_{in}} = \frac{\frac{g_m R_K}{1 + j\omega R_K C_o}}{\frac{g_m R_K}{1 + j\omega R_K C_o} + 1 + j\omega C_s R_{in}} \\ = g_m R_K \frac{1}{g_m R_K + 1 - \omega^2 C_s C_o R_{in} R_K + j\omega(C_s R_{in} + C_o R_K)}$$

Neglecting 1 in comparison with $g_m R_K$, we now consider $C'R'$, the geometric mean of $C_o R_K$ and $C_s R_{in}$, and define a so that $C_o R_K = (1/a) C'R'$ and $C_s R_{in} = a C'R'$. Then, letting $C'R' = 1/\omega_o$

$$\left| \frac{V_{out}}{V_{in}} \right| = g_m R_K \frac{1}{\left\{ \left[g_m R_K - \left(\frac{\omega}{\omega_o} \right)^2 \right]^2 + \left[\frac{\omega}{\omega_o} \left(a + \frac{1}{a} \right) \right]^2 \right\}^{1/2}}$$

This has a maximum near the frequency where $(\omega/\omega_o)^2 = g_m R_K$, i.e. where

$$\omega = \left(\frac{g_m}{C_o C_s R_{in}} \right)^{1/2}$$

THE CATHODALLY SCREENED CATHODE FOLLOWER

and

$$\left| \frac{V_{out}}{V_{in}} \right| = \frac{(g_m R_K)^{1/2}}{a + 1/a}$$

$a + 1/a$ has a rather blunt minimum value of 2 when $a = 1$. Thus, if the two time constants $C_o R_K$ and $C_s R_{in}$ are approximately equal, the gain rises to $(g_m R_K)^{1/2} / 2$. $g_m R_K$ is typically 100, so the gain will rise to 5.

Some curves for this circuit for $g_m R_K = 100$ and various a are plotted in Graph 33. Evidently to avoid the hump in the characteristic a must be >1 or <1 , but it does not much matter which; that is, $C_o R_K$ and $C_s R_{in}$ must be dissimilar. The turn-over frequency of the device, for a given valve and electrode, is inversely proportional to the product of C_o and C_s .

Transient response

For the analysis which follows I am indebted to Professor A. L. Hodgkin. In Figure 18.4

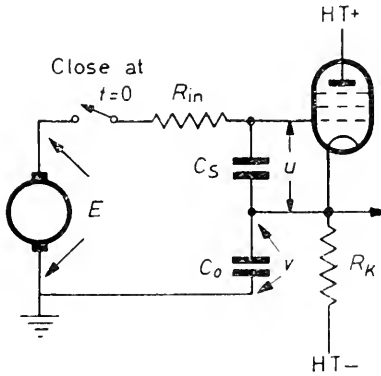


Figure 18.4

If u and v are departures from the initial voltages,

$$E = v + u + R_{in} C_s \frac{du}{dt} \quad \dots(1)$$

$$C_s \frac{du}{dt} + g_m u = C_o \frac{dv}{dt} + \frac{v}{R_K} \quad \dots(2)$$

Combining

$$(R_{in} C_s R_K C_o) \frac{d^2 v}{dt^2} + (R_{in} C_s + R_K C_o + R_K C_s) \frac{dv}{dt} + (g_m R_K + 1)v = g_m R_K E + R_K C_s \frac{dE}{dt}$$

The roots are

$$\frac{-(R_{in} C_s + R_K C_o + R_K C_s) \pm \sqrt{\{(R_{in} C_s + R_K C_o + R_K C_s)^2 - 4 R_{in} C_s R_K C_o (1 + g_m R_K)\}^{1/2}}}{2 R_{in} C_s R_K C_o}$$

At $t = 0$, $v = 0$ and $dv/dt = E/(R_{in} C_o)$. As ordinarily used, $R_{in} \gg R_K$ so

THE CATHODALLY SCREENED CATHODE FOLLOWER

that $R_K C_s \ll R_{in} C_s$. Calling $T_1 = R_{in} C_s$ and $T_2 = R_K C_o$, and $A = g_m R_K$, roots are

$$\frac{-(T_1 + T_2) \pm \{(T_1 + T_2)^2 - 4 T_1 T_2 (1 + A)\}^{1/2}}{2 T_1 T_2}$$

This is maximally oscillatory when $T_1 = T_2$, and reduces to

$$\frac{1}{T} \pm \frac{1}{T} (-A)^{1/2}$$

If A is large (e.g. 100) the solution is approximately

$$v = E \cdot \frac{A}{1 + A} \left\{ 1 - e^{-\frac{t}{T}} \cos \frac{(A)^{1/2}}{T} t \right\}$$

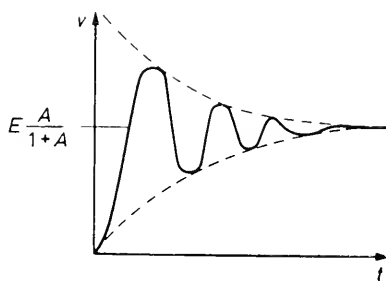


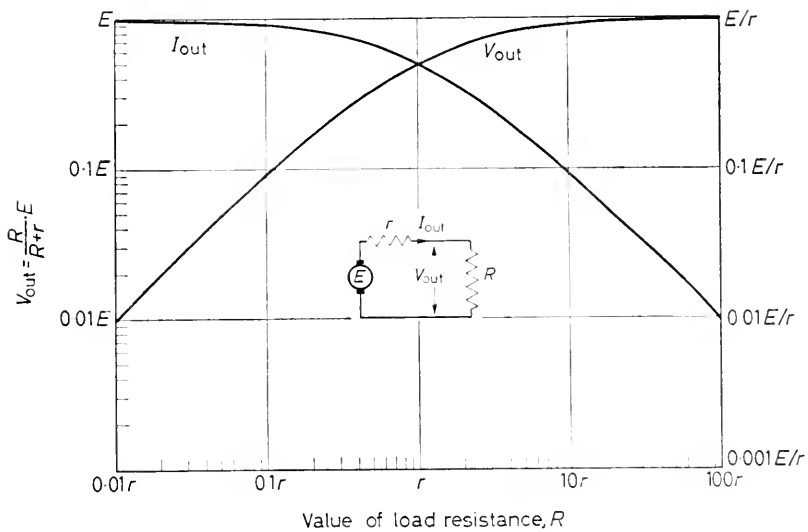
Figure 18.5

The ringing frequency is $(A)^{1/2}/T$

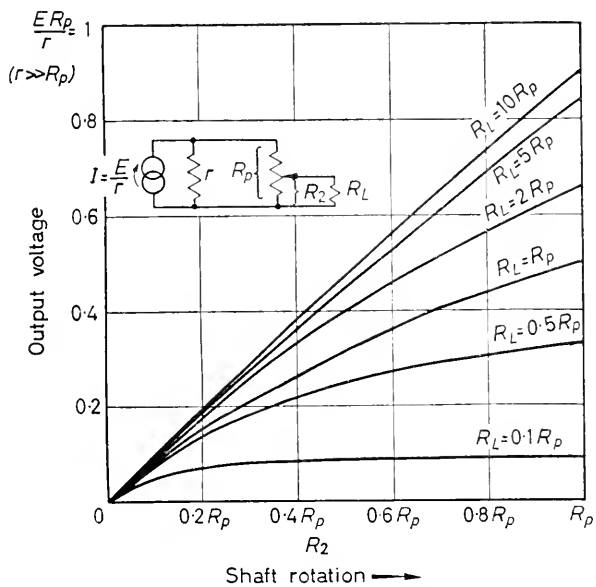
The amplitude declines by $e^{-2\pi/(A)^{1/2}}$ in each cycle.

Readers who care to substitute back the original values into the expression for the ringing frequency will find it is the same as that near which the 'resonance' occurs in the steady-state analysis.

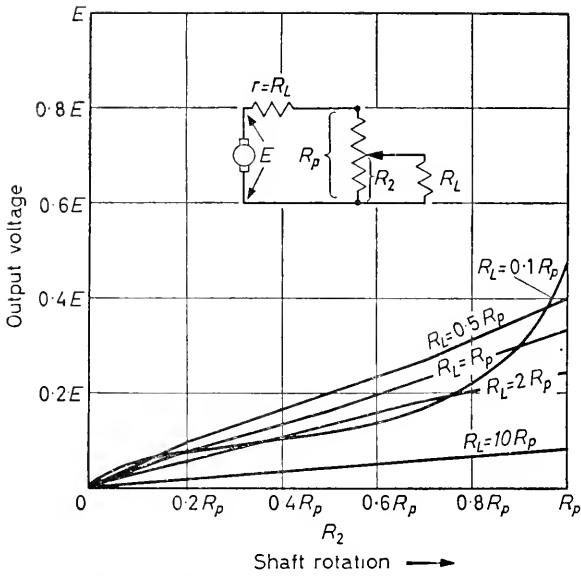
For triode cathode followers the figure for A would be not $g_m R_K$ but $(\mu R_K)/(R_K + r_a)$.



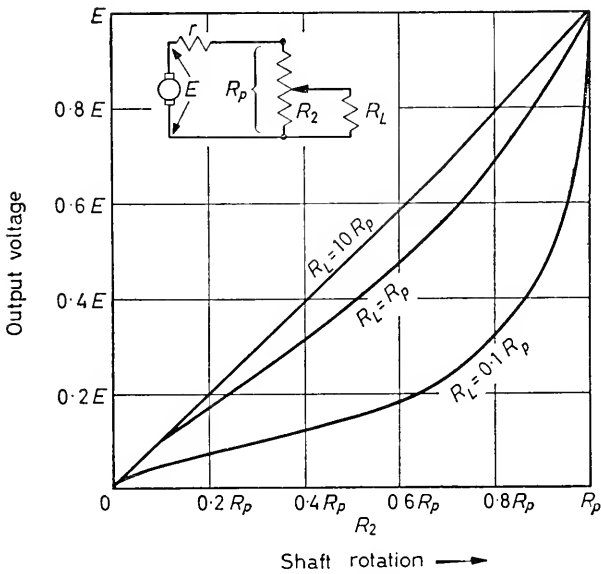
Graph 1



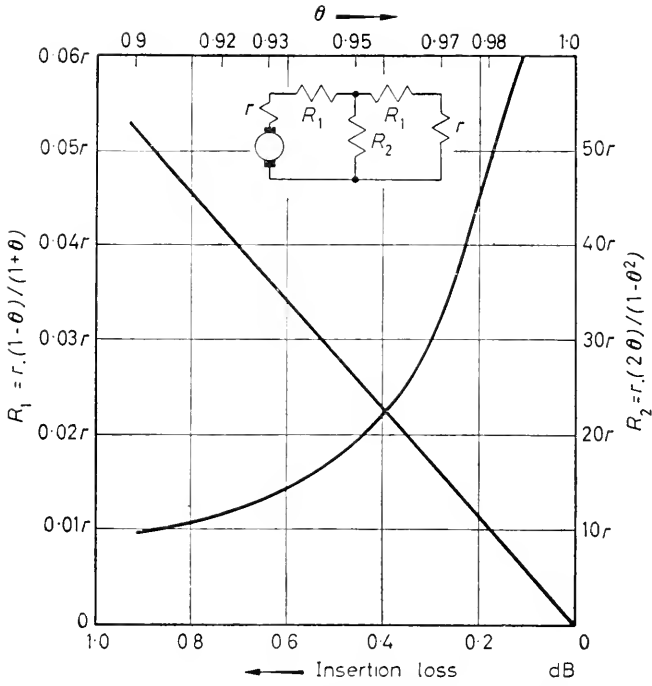
Graph 2. Potentiometric control of power. Potentiometer fed from constant current-type generator



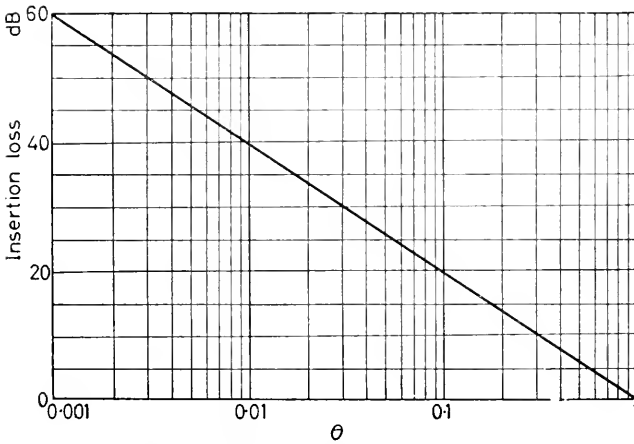
Graph 3. Potentiometric control of power. Potentiometer fed from generator matched to load



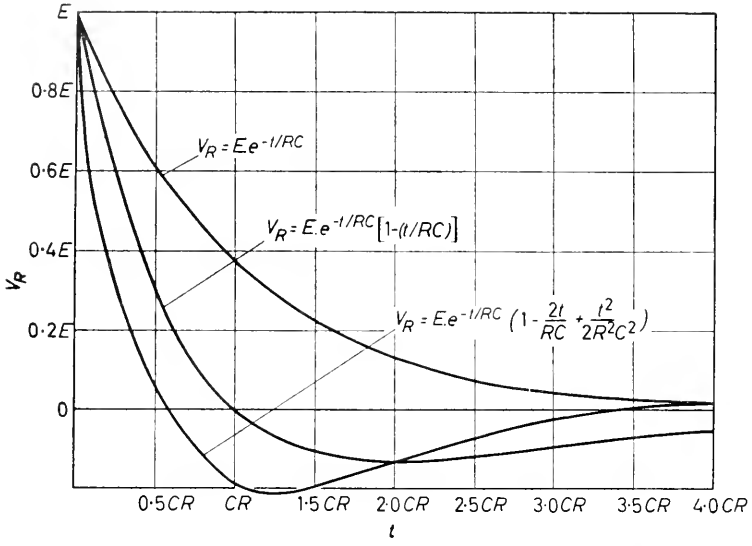
Graph 4. Potentiometric control of power. Potentiometer fed from constant-voltage type generator



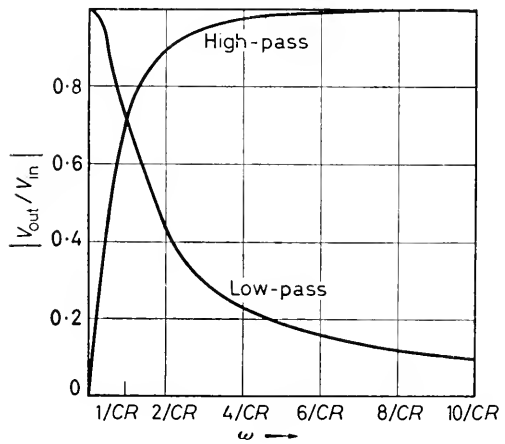
Graph 5. Resistance values in T attenuator section



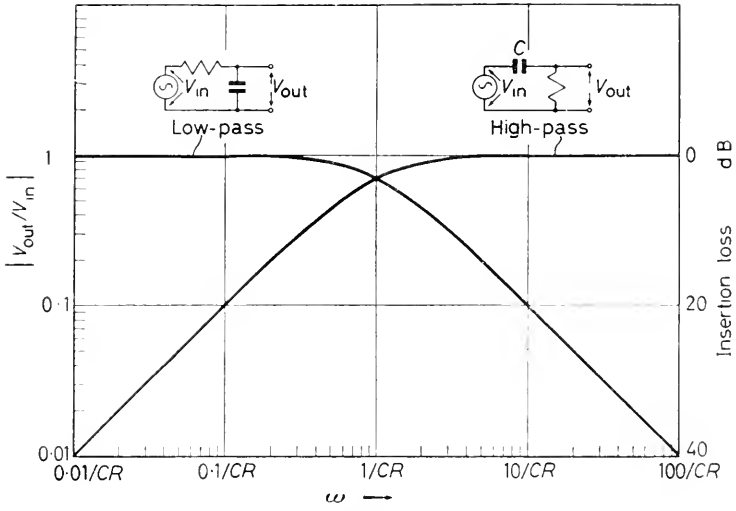
Graph 6. Insertion loss in dB's = $20 \log_{10} 1/\theta$



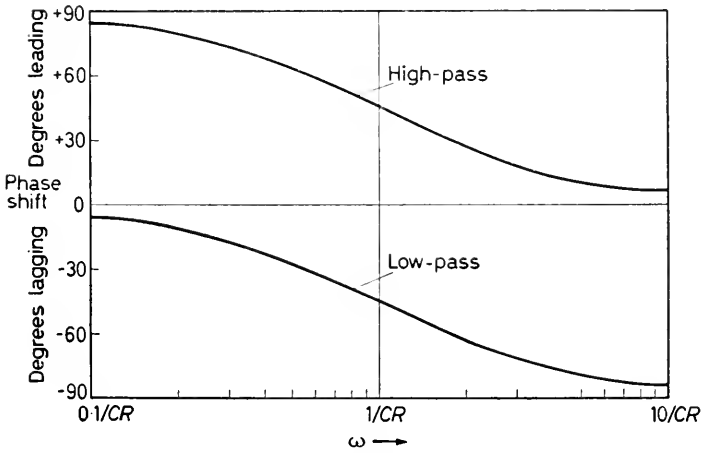
Graph 7. Transient response of 1, 2 and 3 stages of R-C coupling



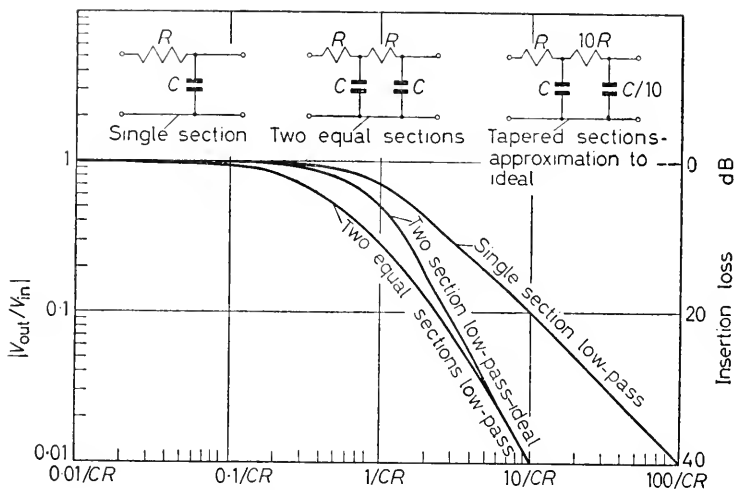
Graph 8. Transmission characteristic of simple R-C filters, on linear scales



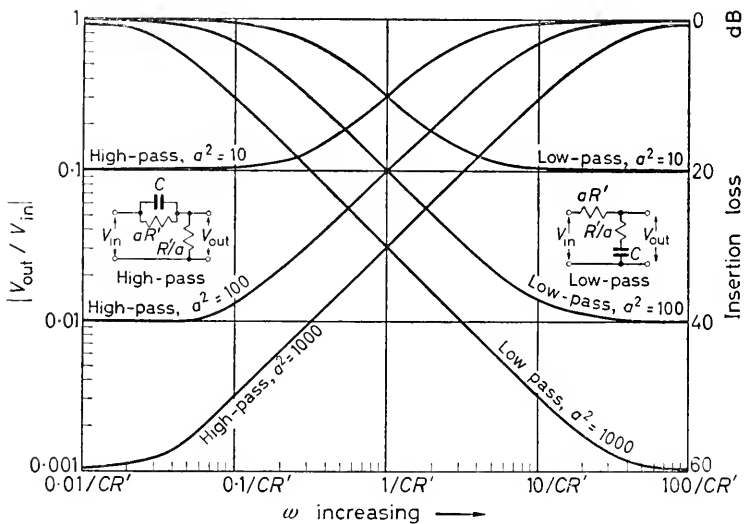
Graph 9. Transmission characteristic of simple R-C filters



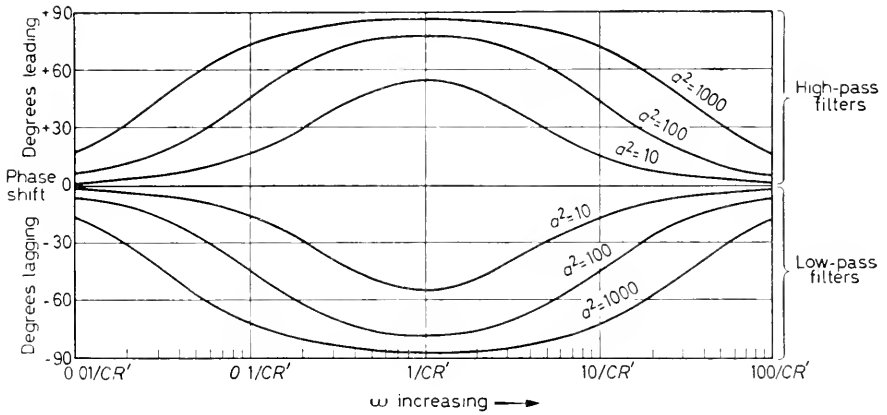
Graph 10. Phase shift of simple R-C filters



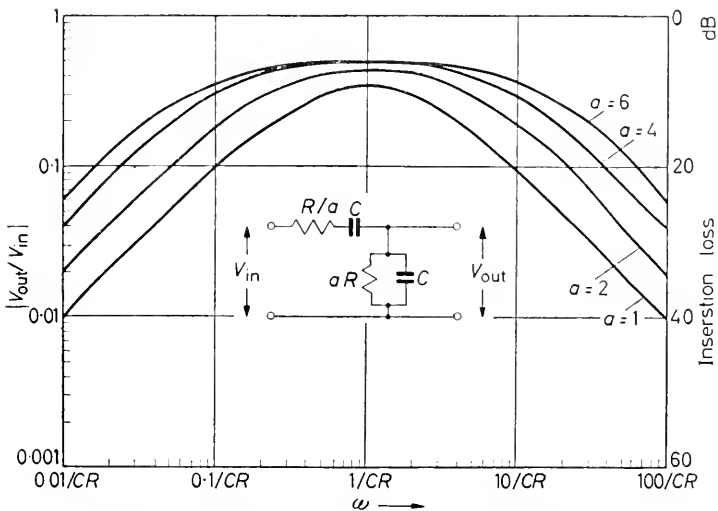
Graph 11. Performance of cascaded R-C sections



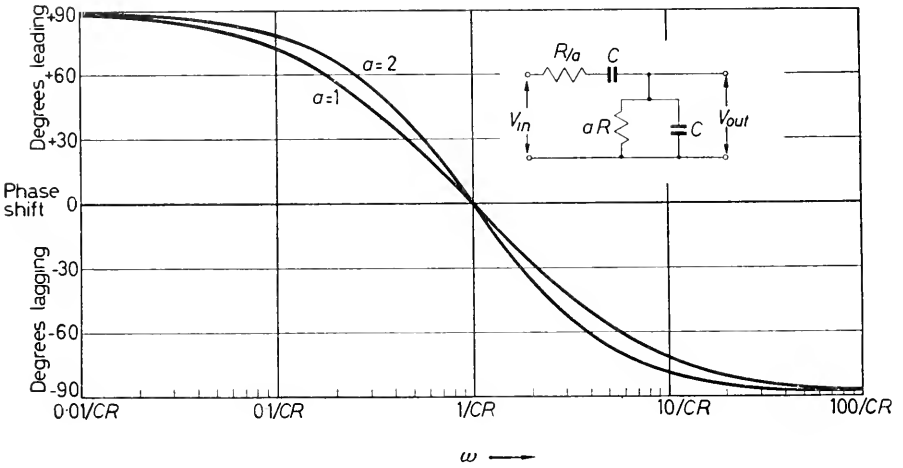
Graph 12. Transmission characteristic of limited phase shift filters



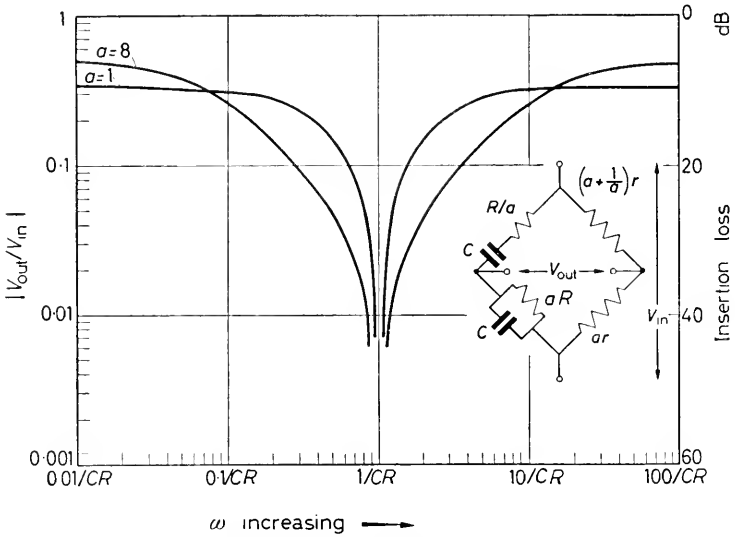
Graph 13. Phase characteristic of limited phase shift filters



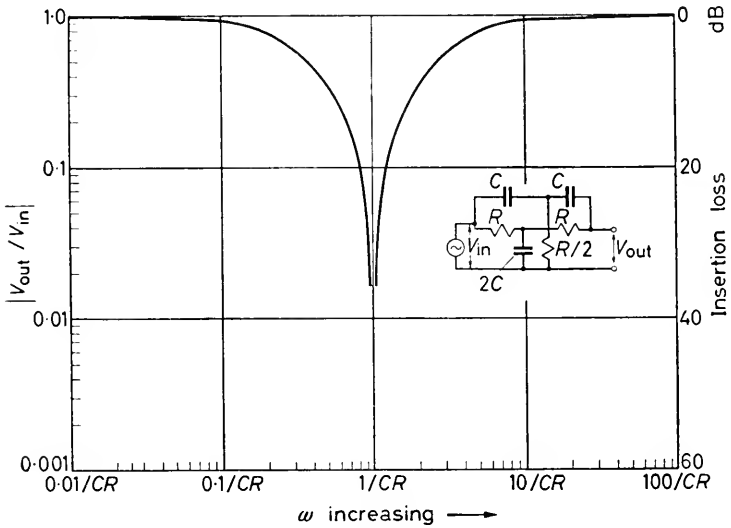
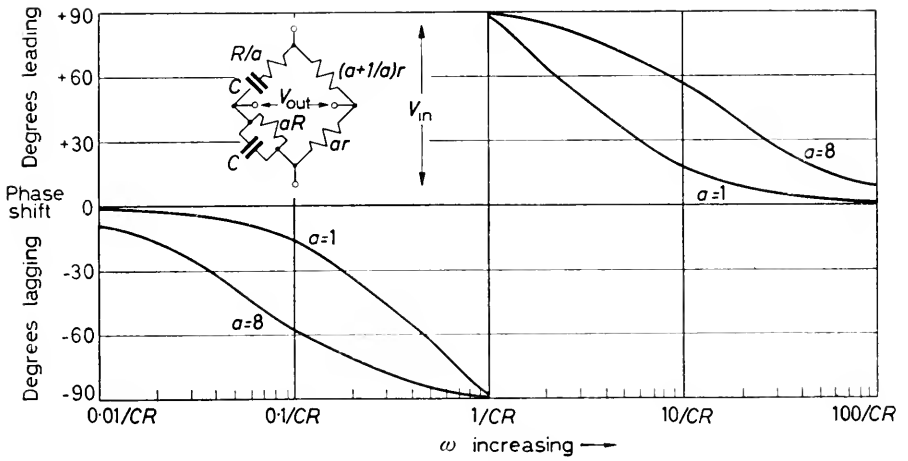
Graph 14. Transmission characteristic of R-C band-pass filter

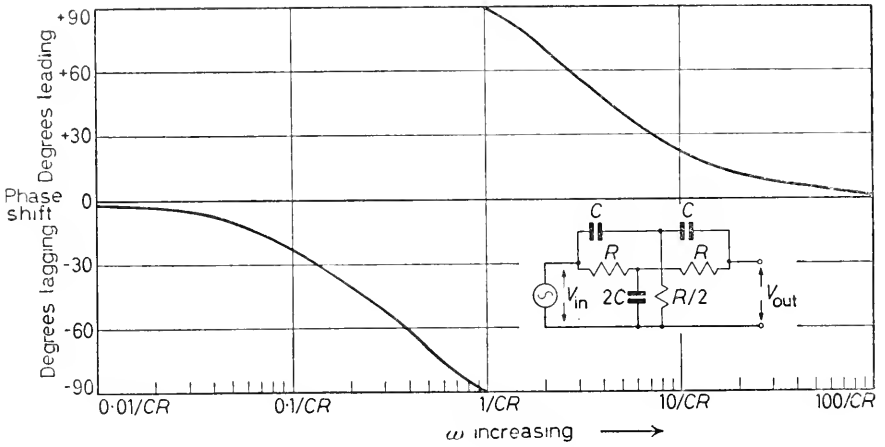


Graph 15. Phase shift of R-C band-pass filter

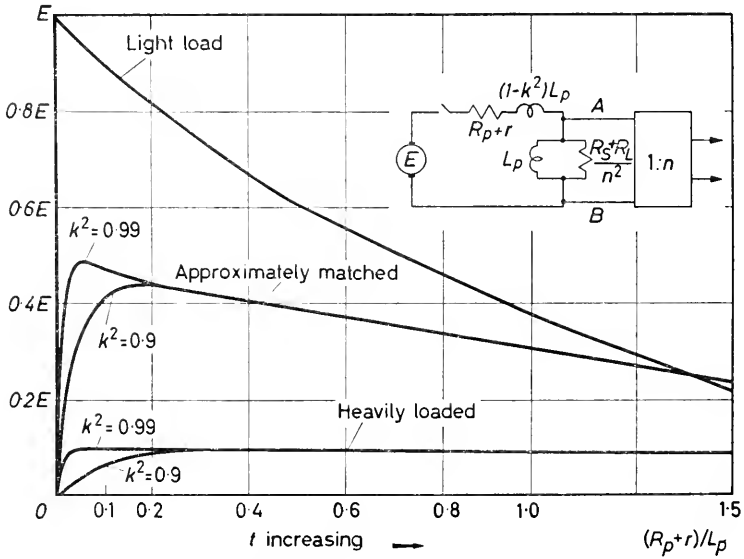


Graph 16. Transmission characteristic of Wien bridge

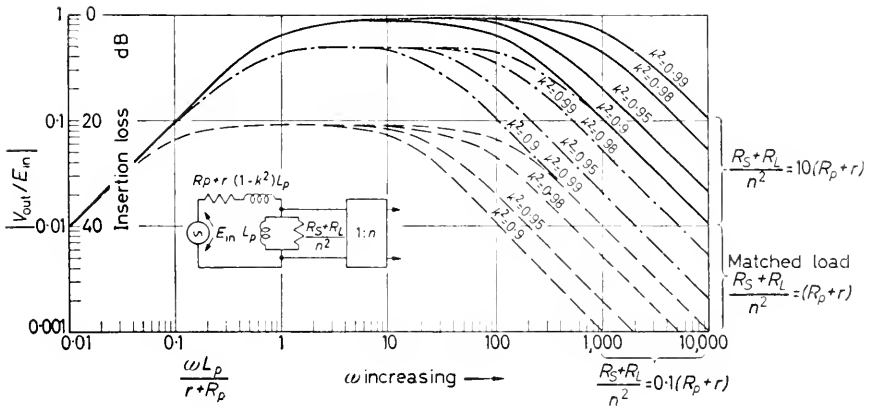




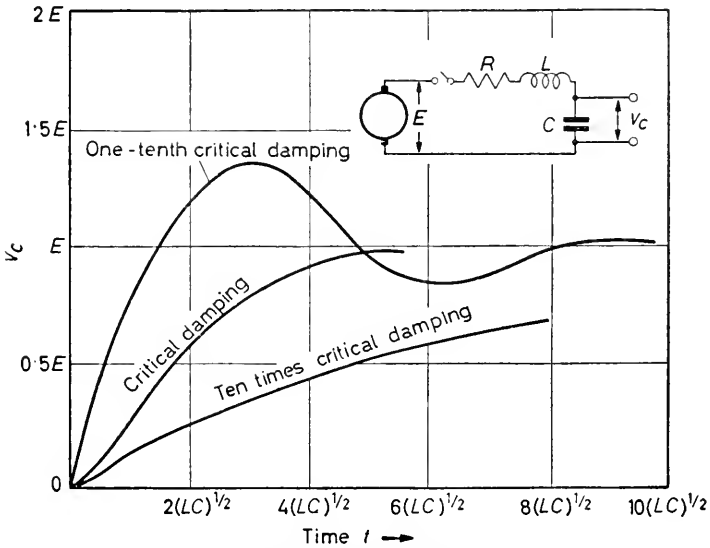
Graph 19. Phase shift of parallel T



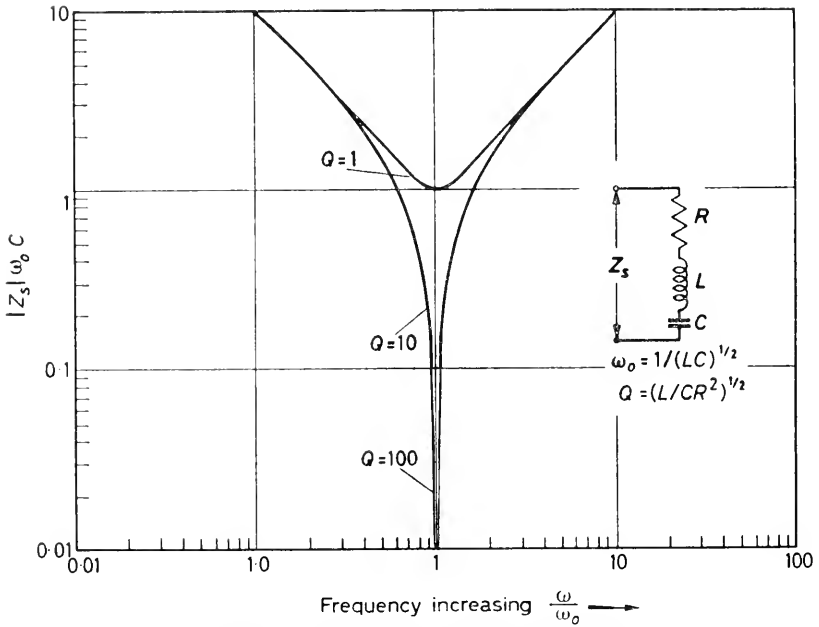
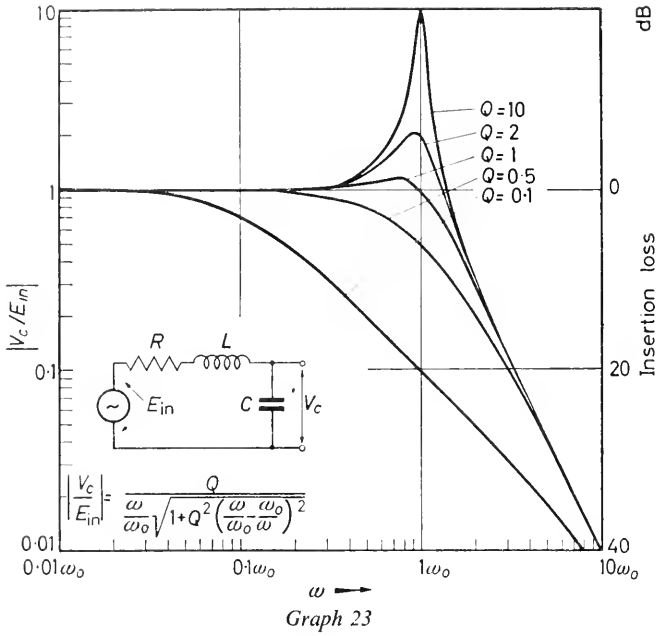
Graph 20. Transient response of pulse transformer

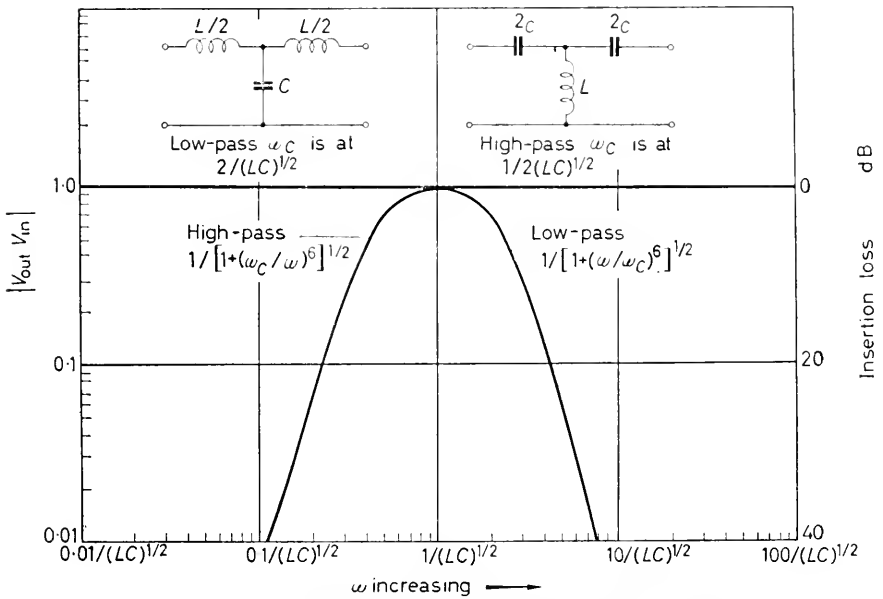
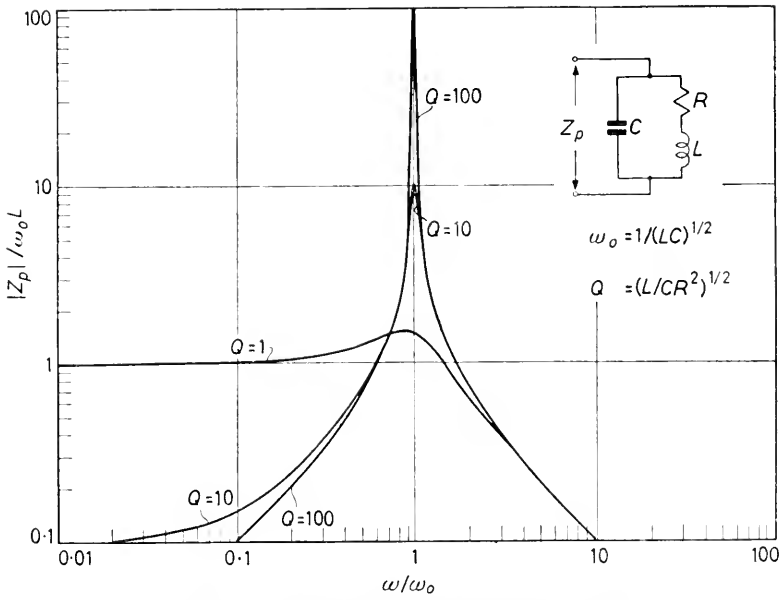


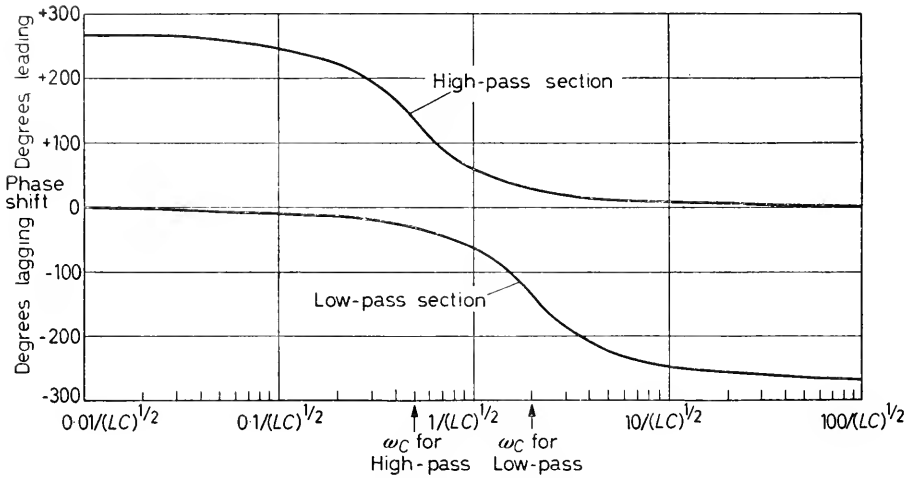
Graph 21



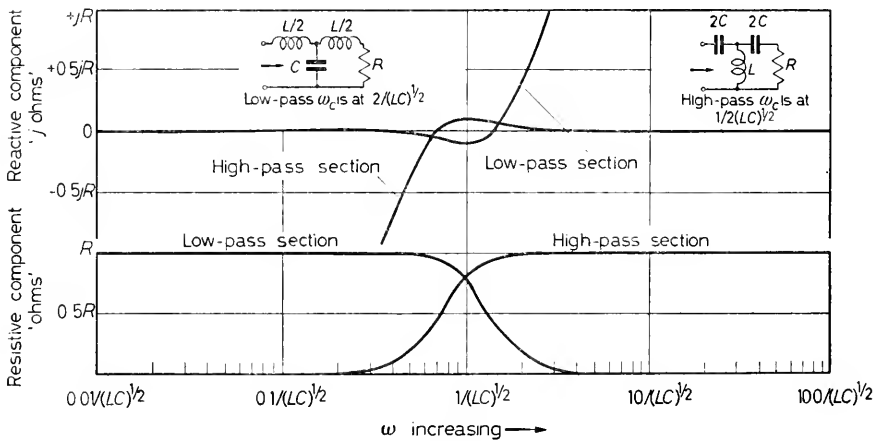
Graph 22. Transient response of series L-C-R circuit



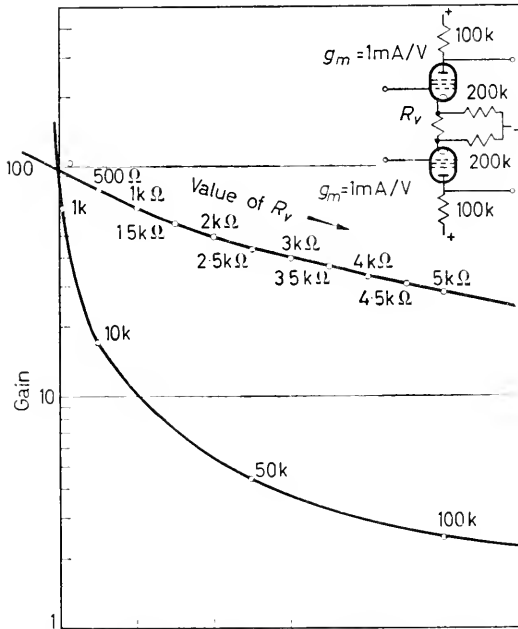




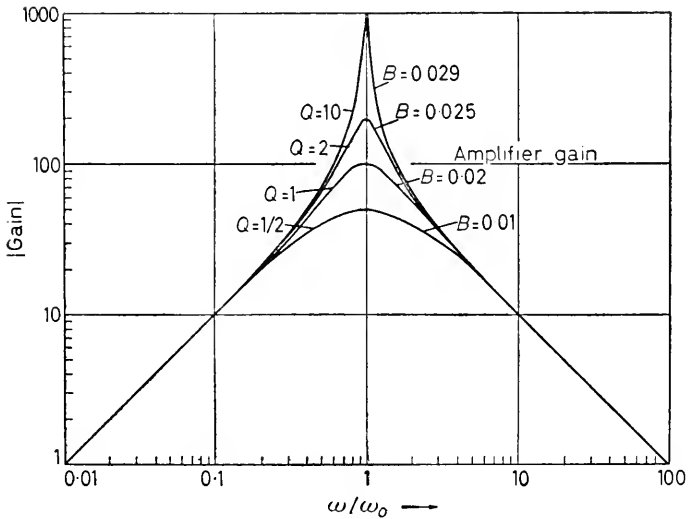
Graph 27. Phase shift of classical filter section



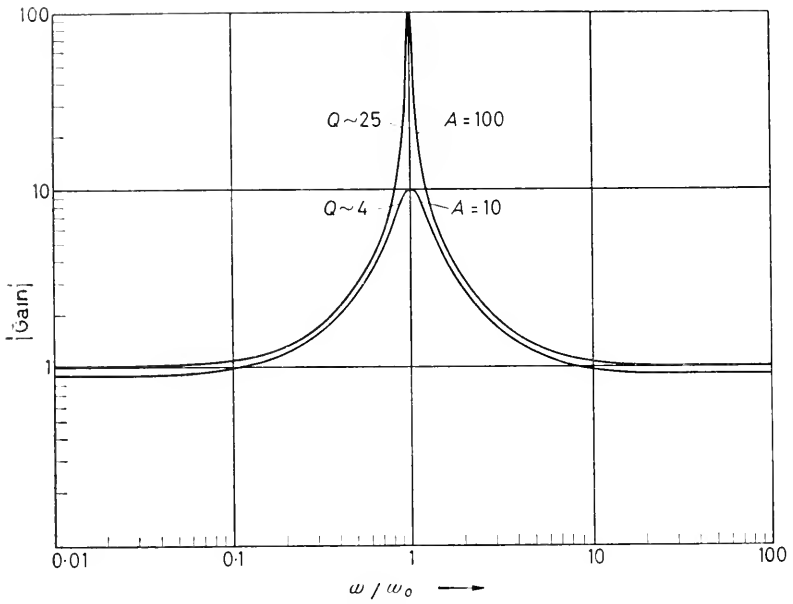
Graph 28. Terminal impedance of classical filter section



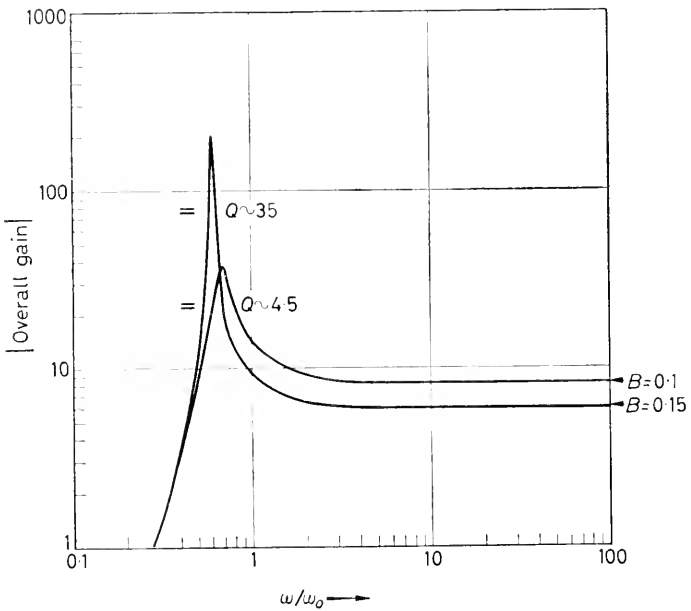
Graph 29. Control law of negative feedback gain-control



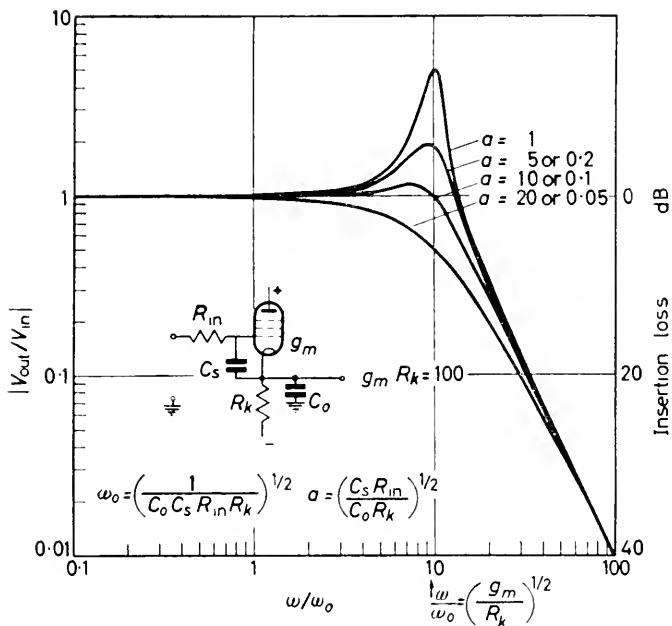
Graph 30. Performance of acceptor-tuned amplifier using the positive feedback method



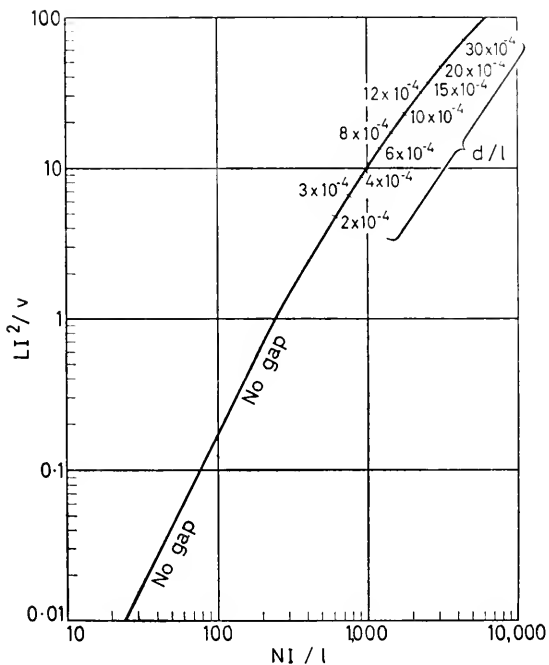
Graph 31. Performance of acceptor-tuned amplifier using the negative feedback method



Graph 32. Performance of acceptor-tuned amplifier—mixed feedback method



Graph 33. Resonance-like phenomenon with a cathode follower



Graph 34. Hanna design curve for A4 Ferroxcube

PART II
PRACTICE

BATTERIES

In portable electrical and electronic apparatus for field use the most generally practicable sources of electric power are the hand, wind, or engine-driven generator and the storage battery. Unless the power supply requirements are particularly heavy the battery is usually the most convenient. In the laboratory there are certain applications for which battery power is more suitable than that derived from the supply mains; these include the supply of power to rather *ad hoc* apparatus whose expectation of life is short, and for which it is not worth building a special power pack; the supply of HT power to the early stages of amplifiers where absence of ripple is essential and where the current demand is so small that replacement costs are unimportant; the supply of extra HT power to photomultiplier cells, for the same reason; the supply of LT (cathode heater) power to the early stages of amplifiers, where it is cheaper to have a simple charger to replenish accumulators between experiments than an elaborate device for making smooth and stable d.c. directly from the mains; where a floating supply is needed having a low capacitance to earth (d.c. amplifier balancing battery, *Figure 12.36* and d.c. pentode cathode follower, *Figure 11.16*); last but not least, as a source of emergency power in experimental runs which would be ruined by a mains supply failure at a crucial moment.

We now discuss some characteristics of the commoner batteries.

PRIMARY BATTERIES

Dry Leclanché—the ordinary 'dry batteries'

These are characterized by ready availability, cleanliness, and a wide range of shapes, sizes and voltages, but their electrical performances do not seem to be widely known. The figures for current outputs contained in the tables below were derived from information supplied to the writer by the courtesy of the Ever Ready Company (Great Britain) Ltd.

Dry Leclanché cells are commonly made in two forms, the traditional cylindrical type in which the electrolyte and depolarizing ingredients are contained in a zinc cup—the negative terminal—whilst the positive terminal is a central carbon rod; and the newer 'layer' type, reminiscent of the original voltaic pile, in which the cells are flat slabs comprising a 3-deck sandwich as in *Figure 19.1*.

Cylindrical cells are used where the discharge current is relatively high, and layer cells when the current is comparatively low. Thus the former are used for LT supplies, flash lamp bulb operation, etc., also HT batteries of the older pattern. Layer type batteries are intended for HT supplies to battery radios, hearing aids, and electronic equipment.

Though there are an enormous number of types of dry battery, they are

BATTERIES

all made up by appropriate series and parallel connections of relatively few types of cell. There are some two dozen common types of cylindrical cells and about the same number of layer cells. We shall not give performance

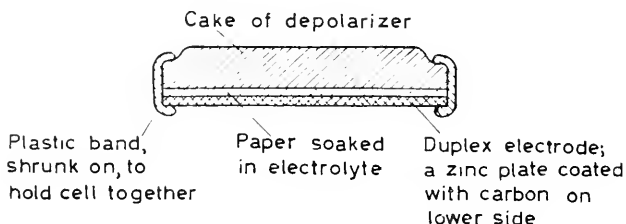


Figure 19.1

figures for all the cylindrical cells but content ourselves with one of the largest (Bell cell), the smallest (D21, hearing aid cell) and two popular intermediate sizes, the U2 and the U10. The latter size of cell is used in the older pattern HT and grid bias batteries—very useful in laboratory work because of the choice of voltages provided by the tapping points*. The cylindrical cells are shown in *Plate 19.1* and the layer cells in *Plate 19.2*.

Dry cells—Initial current to endpoint of 1.1 V.

Duty ratio—4 hours in 24.

Cell	Approximate dimensions in.		Approx. weight oz.	Approximate initial mA at			
				1,000 h rate	100 h rate	10 h rate	
112	1	$\frac{9}{16}$	0.12	0.06	—	0.87	3.9
132	$1\frac{1}{4}$	$\frac{7}{8}$	0.13	0.15	0.27	2.2	10
135	$1\frac{1}{4}$	$\frac{7}{8}$	0.22	0.25	0.35	4.2	18
155	$2\frac{1}{8}$	$1\frac{1}{2}$	0.32	1.1	1.6	18	65
D21	$\frac{43}{32}$	Dia.	$\frac{7}{8}$	0.13	—	1.7	6.8
U10	$\frac{43}{16}$	Dia.	$2\frac{5}{16}$	1.1	1.8	12	55
U2	$1\frac{11}{16}$	Dia.	$2\frac{13}{16}$	3	3.2	39	130
Flag	$2\frac{9}{16}$	Dia.	$6\frac{3}{8}$	32	45	300	—

With this information we can construct two further tables. Taking the average voltage per cell on discharge to be 1.3.

Approximate Energy Delivered per Unit Volume (Wh/in³)

Cell	1,000 h rate	100 h rate	10 hr rate
112	—	1.44	0.65
132	2.14	1.75	0.79
135	1.64	1.96	0.84
155	1.76	1.99	0.72
D21	—	1.70	0.68
U10	1.69	1.13	0.52
U2	1.06	1.29	0.42
Flag	1.48	0.99	—

* We also consider 4 important sizes of larger cell, designated by the Ever Ready Company types 112, 132, 135 and 155.

PRIMARY BATTERIES

Approximate Energy Delivered per Unit Weight (Wh/oz.)

Cell	1,000 h	100 h	10 h
112	—	1.63	0.73
132	2.02	1.65	0.75
135	1.58	1.89	0.81
155	1.64	1.84	0.67
D21	—	1.47	0.59
U10	1.84	1.23	0.56
U2	1.20	1.47	0.49
Flag	1.59	1.06	—

If the endpoint can be reduced to 0.9 V per cell, the life of dry Leclanchés is usually appreciably increased; if the discharge rate is relatively high it may even be doubled.

Reactivation of dry cells

This subject has received some attention in recent years^{1,2} and may therefore be of interest.

If batteries are exhausted by discharge, re-charging, or 'reactivation' as it is commonly called, is not practicable. Under carefully controlled cycles of charge and discharge with appropriate current consumption a measure of reactivation is possible, but to obtain satisfactory results close control is essential, and whilst attainable in the laboratory it cannot usually be effected by the average user. Moreover if the proper conditions of reactivation are not accurately maintained, battery life may be shortened rather than lengthened.

Vidor 'Kalium' cell

These are generally similar to the cylindrical dry Leclanché cell, but employ mercuric oxide as depolarizer. They are stated to have a better shelf-life, particularly for apparatus to be used under high temperature conditions. It has been reported³ that they make excellent reference sources of e.m.f., having a lower temperature coefficient (+17 parts per million per degree Centigrade) even than the Weston's standard cell.

*Mallory cells**

In these cells the negative pole is zinc, the positive pole is mercuric oxide and graphite and the electrolyte is potassium hydroxide. The whole is enclosed in a steel case, giving good dimensional stability. In general, unlike the dry Leclanché, the case is the positive pole. These cells are characterized by an extremely constant voltage of about 1.3 during most of the discharge period, very high relative discharge rates, good ratios of capacity to weight and volume, and rather high cost. The makers also claim that the energy output is independent of duty ratio.

From the two dozen or so types of cell at present manufactured figures are given for a large one (type RM 12), and intermediate size (type RM 1) and a small one (type TR 112). Type RM 1 is shown in *Plate 19.3*.

* Mallory Batteries Ltd., Dagenham, Essex.

BATTERIES

The following is an extract from the makers' pamphlet:

<i>Type</i>	<i>Capacity mAh</i>	<i>Nominal voltage</i>	<i>Diameter in.</i>	<i>Height in.</i>	<i>Weight oz.</i>	<i>Recommended max. current drain (mA)</i>
TR 112	250	2.6 (2 cell unit)	0.65	0.6	0.303	20
RM 1	1,000	1.34	0.62	0.65	0.43	100
RM 12	3,600	1.34	0.62	1.95	1.6	250

It follows from the above that:

<i>Type</i>	<i>Energy delivered per unit volume Wh/in.³</i>	<i>Energy delivered per unit weight Wh/oz.</i>
TR 112	3.27	2.15
RM 1	6.87	3.12
RM 12	8.25	3.02

The Leclanché and the Mallory type cells should be regarded as complementary. Where very small size and volume are of critical importance the latter have obvious advantages. Similarly when small size and volume are not a first consideration the Leclanché dry cells are more suitable and more economic. A comparison of cost in terms of pence per Wh is given in the table below:

<i>Battery</i>	<i>Contents</i>	<i>Cost, pence</i>	<i>Approx. cost/cell, pence</i>	<i>Approx. cost/ unit energy (pence per Wh at 100 h rate)</i>
B123	20-112	39	2	20.4
B119	20-132	51	2½	10.1
B126	60-135	120	2	4.2
B107	60-155	177	3	1.5
D21	—	5	5	26.0
U10	—	4½	4½	3.4
U2	—	7	7	1.6
Flag	—	66	66	2.0
TR 112	—	54	54	83
RM 1	—	27	27	20
RM 12	—	57	57	11.7

These figures show that, output for output, the dry Leclanché cell is a good deal cheaper.

PRIMARY BATTERIES

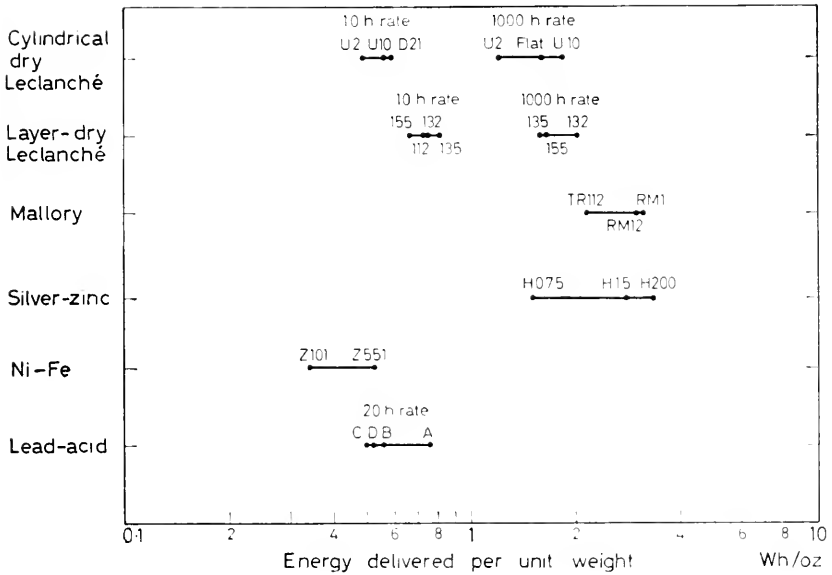


Figure 19.2

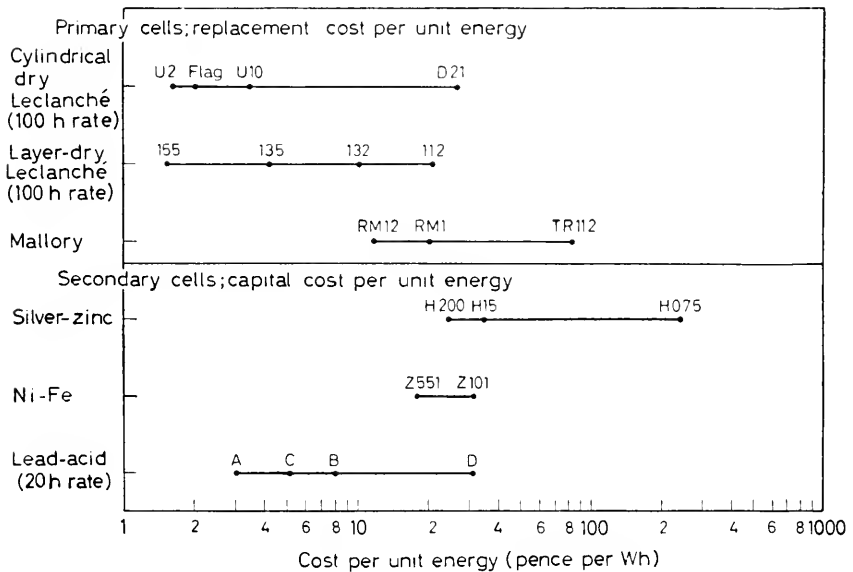


Figure 19.3

BATTERIES

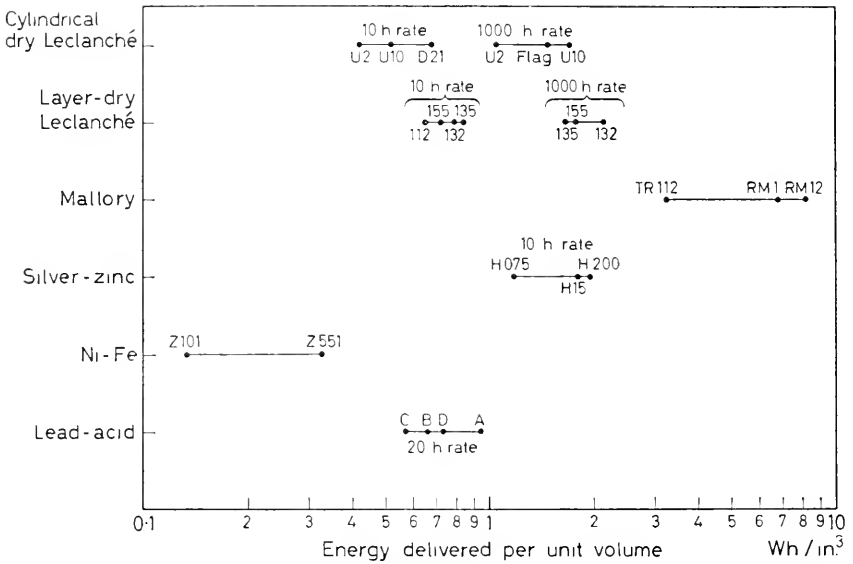


Figure 19.4

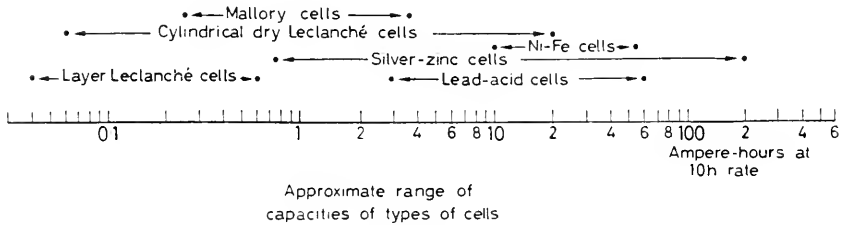


Figure 19.5

SECONDARY CELLS

The lead-acid cell or 'accumulator'

These are too well known to require description. They have a steady voltage of 2 during the greater part of the discharge period. On first connecting the load to a freshly charged cell the voltage may be 2.2, and discharge should be stopped when the voltage has fallen to 1.8. Accumulators can be quickly damaged if proper attention is not paid to their maintenance, and at best their life is about 5 years. To investigate their storage properties we take four examples, all of which might be found supplying electronic equipment:

Battery A—a 12 V car battery

Battery B—a 6 V motor cycle battery

Battery C—a glass encased 'laboratory' type of cell

Battery D—a small celluloid-encased cell.

These are shown in *Plate 19.4*, and details are:

SECONDARY CELLS

Battery	Voltage	Capacity at 20 h rate (Ah)	Capital cost	Weight	Dimensions (in.)
A	12	60	£10 10 0	59 lb.	$12\frac{1}{2} \times 6\frac{13}{16} \times 9\frac{1}{8}$
B	6	12	£2 4 0	128 oz.	$4\frac{13}{16} \times 3\frac{5}{8} \times 6\frac{3}{8}$
C	2	40	£1 14 0	160 oz.	$8 \times 5 \times 3\frac{1}{2}$
D	2	3	15 6d	11.4 oz.	$3\frac{1}{8} \times 3 \times \frac{7}{8}$

From the above we can construct the following:

Battery	Energy delivered per unit volume (VAh/in. ³)	Energy delivered per unit weight (VAh/oz.)	Current for 20 h discharge (A)	Capital cost per energy delivered (pence/VAh)
A	0.936	0.760	3	3
B	0.657	0.563	0.6	7.85
C	0.572	0.50	2	5.1
D	0.735	0.525	0.15	31

*Nickel-iron-cadmium cells or 'Ni-Fe' cells**

These are steel-cased cells whose electrodes are steel plates in the form of a matrix of perforated pockets. These pockets contain, for the negative plates, a preparation of cadmium and iron, and for the positive plates, nickel hydroxide. The electrolyte is an aqueous solution of hydroxide. The voltage of discharge at normal rates is 1.2. Advantages of the Ni-Fe cell over the lead acid cell are absence of corrosive fumes and spray (an important point with delicate apparatus), a very long life (15-20 years), virtual indestructibility—both electrical and mechanical (the cells may be overcharged, left uncharged, short-circuited, all without harm)—and ability to hold charge for long periods. These points should be borne in mind when looking at the performance figures which show the Ni-Fe cell in a somewhat unfavourable light compared with the other secondary cells. We take two examples from the manufacturers' catalogue of portable batteries mounted in hardwood crates for general laboratory use.

Type	Capacity (Ah)	V	Weight (oz.)	Dimensions (in.)	Cost
Z 101	10	1.2	35.2	$2\frac{13}{16} \times 4\frac{7}{16} \times 7\frac{1}{4}$	32/8d
Z 551	55	1.2	125	$4\frac{1}{8} \times 5\frac{3}{8} \times 9\frac{1}{2}$	98/1d

* NIFE Batteries, Redditch.

BATTERIES

hence

<i>Type</i>	<i>Energy delivered per unit volume (VAh/in.³)</i>	<i>Energy delivered per unit weight (VAh/oz.)</i>	<i>Capital cost/unit energy (pence/VAh)</i>
Z 101	0.133	0.34	32.6
Z 551	0.313	0.527	17.8

It is only fair to mention that the makers produce cells of special design for certain electronic applications, and these show a substantial saving in weight and volume.

*Silver-zinc cell**

The outstanding characteristic of these cells is their energy storage per unit volume and weight, as the figures will show. Unfortunately the cost is rather high. The cell voltage is 1.5 for the greater part of the discharge. The electrolyte is potassium hydroxide. We take three examples from the makers' catalogue.

<i>Type</i>	<i>Capacity (Ah)</i>	<i>V</i>	<i>Weight</i>	<i>Dimensions (in.)</i>	<i>Cost</i>
H075	0.75	1.5	$\frac{3}{4}$ oz.	$1\frac{1}{2} \times 1\frac{1}{8} \times \frac{9}{16}$	£1 2 6
H15	15	1.5	8 oz.	$4\frac{13}{16} \times 2\frac{5}{16} \times \frac{3}{4}$	£3 5 0
H200	200	1.5	89 oz.	$8 \times 4 \times 3\frac{13}{64}$	£30 10 0

hence:

<i>Type</i>	<i>Energy stored per unit volume (10 h rate) (VAh/in.³)</i>	<i>Energy stored/unit weight (VAh/oz.) (10h rate)</i>	<i>Capital cost/unit energy (pence/VAh)</i>
H075	1.19	1.5	240
H15	1.8	2.82	34.7
H200	1.96	3.36	24.4

In the accompanying illustrations an attempt has been made to collect and display the significance of the figures in the foregoing tables.

REFERENCES

- ¹ HALLOWS, R. W. *Wireless World* 59 (1953) 344
- ² — *Wireless World* 61 (1955) 503
- ³ JERVIS, M. W. *J. sci. Instrum.* 30 (1953) 60

* Venner Accumulators Ltd., New Malden, Surrey.

RESISTORS

FIXED RESISTORS

In assessing the suitability of a fixed resistor for a particular application in electronics it is necessary to pay attention to a number of factors beside the value in ohms. These are:

(1) The material of which the resistor is made: the possibilities are carbon composition, cracked carbon, or wirewound.

(2) The power rating. This is the maximum electrical power which may be dissipated in the component as heat in order that the consequent temperature rise shall not damage it. Strictly speaking the rating has not much meaning unless the ambient temperature is also taken into account. Thus a component sold as a 'half-watt resistor' should not be expected to dissipate half a watt immediately above a 25 watt power pentode in miniaturized equipment for tropical use.

(3) The voltage rating. This is the maximum potential difference which may be allowed to occur across the resistor in order that the electrical field within the component shall not damage it.

Carbon composition resistors

These are much the commonest resistors in electronic practice; despite their defects they are cheap, extremely reliable if properly used and perfectly satisfactory for the general run of electronic circuitry. The resistive element is a stick composed of a compressed mixture of carbon black and a resin binder. In the 'uninsulated' variety connection is made to the stick by tinning the ends, wrapping round the connecting wires, and soldering the join to make good contact; the wires leave the resistor perpendicularly (*Plate 20.1*). In the 'insulated' kind the element is encased in a ceramic tube, and the connecting wires enter axially at the ends (*Plate 20.2*). Modern resistors tend to be of the insulated type: an advantage of the uninsulated kind is the facility for adjusting the value of resistance upward by filing the component thinner.

Resistance values of composition resistors are marked by means of a colour code. The meaning of the colours is as follows:

black	zero	blue	six
brown	one	violet	seven
red	two	grey	eight
orange	three	white	nine
yellow	four	gold	$\pm 5\%$ tolerance
green	five	silver	$\pm 10\%$ tolerance

Insulated resistors are marked by bands of colour which completely encircle them, so that the value may be read from any aspect (*Figure 20.1*). The first band, identifiable by being nearer one end than the fourth band is to

RESISTORS

the other, gives the first digit of the resistance, the second band the second digit, the third band the number of zeros and the fourth band the tolerance. If the fourth band is absent it means the tolerance is ± 20 per cent.

Uninsulated resistors use the 'body-tip-spot' system (*Figure 20.2*), where the body colour gives the first digit, the tip colour the second and the spot

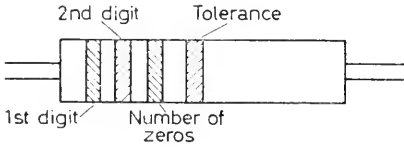


Figure 20.1

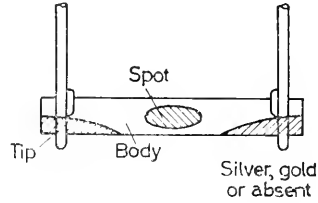


Figure 20.2

the number of zeros. There may be a tip colour—silver or gold—at the other end which indicates the tolerance; if there is none, the tolerance is ± 20 per cent.

In the case of resistors between 1 and 9 ohms inclusive the values—so far as applying the code is concerned—are 01 to 09. Thus a 6 ohm uninsulated resistor ± 20 per cent has a first digit zero (black body), second digit 6 (blue tip) no zeros (black spot) no tolerance mark (other tip unmarked). The component is therefore entirely black except for one blue tip.

Newcomers to electronics are occasionally puzzled by the peculiar values of resistances which seem to be common, 47 k, 68 k, etc. These numbers follow from the use of the 'preferred value' system, in which the nominal resistor values which should be manufactured are not arbitrary, but are dictated by the tolerance required. This is explained in *Figure 20.3* which

1										1.5					2.2					3.3					4.7					6.8					10					$\pm 20\%$														
1	1.2	1.5	1.8	2.2	2.7	3.3	3.9	4.7	5.6	6.8	8.2	10	$\pm 10\%$																																									
1	1.1	1.2	1.3	1.5	1.6	1.8	2.0	2.2	2.4	2.7	3.0	3.3	3.6	3.9	4.3	4.7	5.1	5.6	6.2	6.8	7.5	8.2	9.1	10	$\pm 5\%$																													
1										2					3					4					5					6					7					8					9					10				

Figure 20.3 Table of preferred values

shows how the 'preferred values' in each tolerance range are cell centre values for a range of adjacent cells of approximately equal width. The actual resistance of a component whose nominal value is the cell-centre value will lie somewhere within that cell. In this way *any* resistor manufactured could be sold under one or other of the nominal values—none need be wasted. In fact composition resistors can be manufactured to within ± 20 per cent of the required nominal value, and the ± 10 and ± 5 per cent specimens are then obtained by measurement and selection. *Figure 20.3* enables one to choose the nearest preferred value to a calculated required one.

Carbon composition resistors suffer from two disadvantages, their instability and their noise. The instability takes the form of a non-reversibility in the resistance-temperature characteristic; as the component undergoes



Plate 19.1 Group of 'cylindrical' type dry Leclanché cells



Plate 19.2 Group of 'layer' type dry Leclanché cells

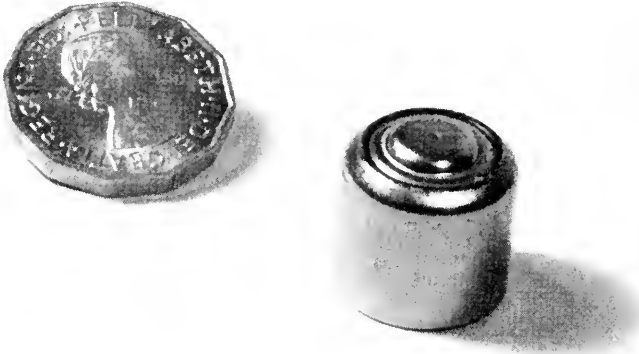


Plate 19.3 Mallory type RMI. This small cell has a capacity of 1 Ah

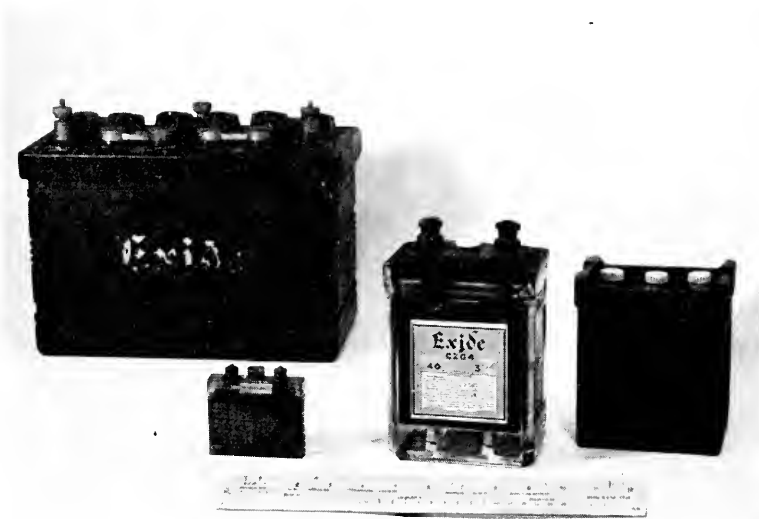


Plate 19.4 Group of lead-acid cells and batteries



Plate 20.1 Uninsulated composition resistors, rated at 3, 2, $1\frac{1}{2}$, 1, $\frac{1}{2}$ and $\frac{1}{8}$ W

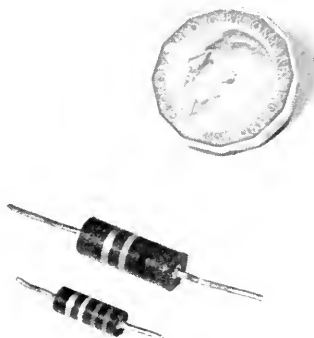


Plate 20.2 Insulated composition resistors, rated at $\frac{1}{2}$ and 1 W



Plate 20.3 Pyrolytic resistors, rated at 2, 1, $\frac{1}{2}$ and $\frac{1}{4}$ W



Plate 20.4 General purpose wirewound resistors, conservatively rated at 15, 10, 6, $4\frac{1}{2}$, 3 and $1\frac{1}{2}$ W. Considerably higher dissipations than these are possible, e.g. 6 W in the 3 W size and 30 W in the 10 W size



Plate 20.5 Precision wirewound resistor

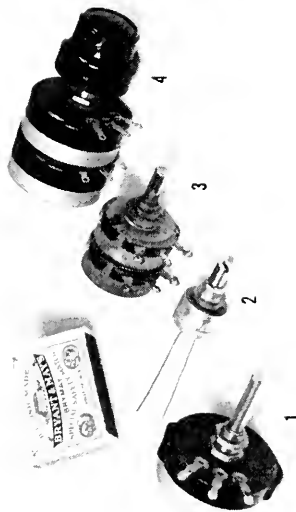


Plate 20.6 Standard, miniature, gauged and tandem potentiometer



Plate 20.7 Plan and cross-section of moulded track



Plate 20.8 Parts of sprayed-track potentiometers. On the left, plastic-based; on the right, fibre-based



Plate 20.9 Interior of linear wirewound potentiometer

FIXED RESISTORS

cyclic temperature changes during the life of the equipment it is in, its resistance is subject to a small but progressive alteration, usually towards a higher value. For this reason it is difficult to attach a temperature coefficient to carbon resistors. The only way to minimize the effect is to choose resistors with an ample margin of safety in power rating.

Composition resistor noise has been briefly referred to in Chapter 17. In addition to Johnson noise, these components develop a noise voltage proportional to any direct current that may be passed through them. This 'current' noise is distributed in frequency in a similar manner to flicker noise in valves, that is, the noise power in a given small band of frequencies is inversely proportional to the mid-band frequency. It follows that current noise is most serious in direct-coupled amplifiers. Its magnitude differs widely from specimen to specimen of a particular type, and nominal value, of resistor. Hence there is no simple expression which may be quoted for giving the reader some idea of the importance of current noise. As a guide, however, an empirical expression given by G.W.A. Dummer (*Fixed Resistors*, Pitman) for the total noise allowed in composition resistors for use by the Armed Services is

$$\bar{V}_{\text{noise}} = 2 + \log_{10} \frac{R}{1,000}$$

microvolts per volts applied across the resistor, where R is the resistance in ohms and the noise is measured over the band 200-10,000 c/s. Current noise is generated to a much greater degree by resistors of low, rather than high, power rating. Thus we have another reason for having an ample margin here.

The voltage rating for composition resistors varies considerably between types. As a rough guide:

Uninsulated—1 watt, 1,000 V	Insulated—1 watt, 500 V
— $\frac{1}{2}$ watt, 500 V	— $\frac{1}{2}$ watt, 250 V

'Cracked carbon', 'high stability', or pyrolytic resistors

In these resistors the resistive element is a piece of ceramic rod on to which a surface film of carbon has been 'cracked' by exposing it to a hydrocarbon vapour at about 1,000°C. Connections are then made to the ends of the rod by pressing on metal caps to which the lead wires are secured, after which the resistor may be adjusted to the proper value by machining away the conducting film along a spiral path until the required value is reached. The whole assembly is then protected by some kind of paint or varnish. In this way components are produced whose resistance is stable, whose temperature coefficient of resistance is less than ± 0.1 per cent per degree centigrade, whose resistance value is generally within 1 per cent of the nominal and whose current noise is about 1-10 per cent of that for a composition resistor. The resistance value in ohms and the tolerance are usually printed on the component. The voltage rating is about 200 for the half-watt size and 400 for the one-watt size. A group of pyrolytic resistors is shown in *Plate 20.3*.

Wirewound resistors

These are of two kinds, general purpose and precision.

RESISTORS

General purpose wirewound resistors—These are used where the power to be dissipated is beyond the capabilities of a carbon component. They comprise a winding of resistance wire (usually nickel-chrome) round a ceramic tube, the winding being secured and protected by a coating of a refractory cement or vitreous enamel. By the use of these fire-resistant materials extremely high working temperatures are possible, and resistors of this type are characterized by remarkably high wattage ratings for their size (*Plate 20.4*). The temperature coefficient is about $+0.015$ per cent per degree C and the voltage rating about 200 V per in. of resistor body length. The tolerance is usually 5 per cent and the stability is comparable with cracked carbon. High value general purpose wirewound resistors (100 k Ω) are often specified for the anode loads in biological amplifiers. These contain a great deal of very thin wire and in the author's experience are very unreliable and are often faulty even before being used. The solution is to use a number of lower values in series.

Precision resistors—These are used in instrument-measuring practice where a very low tolerance, $+0.1$ per cent, and good long-term stability, less than 1 per cent in value, are required. The winding is of nickel-chrome or nickel-copper alloy, and, like the general purpose type, is often wound on a ceramic tube. They are, however, very definitely not of fireproof construction, containing as they do materials such as silk, rubber and paper; they are essentially 'room-temperature' devices and are consequently large for their power ratings. A temperature coefficient of $+0.005$ per cent per degree C may be regarded as typical. A 1-M Ω precision resistor is shown in *Plate 20.5*.

VARIABLE RESISTORS OR POTENTIOMETERS

The only difference between a variable resistor and a potentiometer lies in the provision of a terminal at only one end, or at both ends, of the 'track'. In view of the trifling extra cost of providing one at either end this is always done; in practice a variable resistor and a potentiometer are the same thing—the only difference lies in the external connections.

In this section we are not concerned with the high grade instrument type potentiometer—these are dealt with under Mechanoelectric transducers in Part III—but with the ordinary cheap component used in radio, television and general electronic practice. A group is shown in *Plate 20.6*. No. 1 is a standard type, and No. 2 a miniature. No. 3 is a 'two-gang' component, two variable resistors controlled together by a single shaft. No. 4 is a 'tandem' potentiometer, in which two units are separately adjustable by a pair of concentric knobs. As with fixed resistors they are of two kinds, carbon or wirewound.

Carbon track variable resistors

These are the cheaper variety. They are obtainable in values between 10 ohms and 5 megohms, but the 'preferred value' system is not used; the sequence of values being 1, 2.5, 5, 10 . . . The tolerance is about ± 25 per cent. They may be made in either of two ways: the carbon track may be moulded out of a similar material to that used for making uninsulated fixed resistors, and might in fact be regarded as such a resistor, but horse-shoe shaped

VARIABLE RESISTORS OR POTENTIOMETERS

instead of straight, *Plate 20.7*; alternatively, the track may be made by spraying the resistive material on to an insulating supporting base, e.g. a fibreboard or a plastic (*Plate 20.8*). Power ratings for carbon track variable resistors range from about $1\frac{1}{2}$ watts for the standard size models to 0.1 watt for the tiny components, less than $\frac{3}{4}$ in. in diameter, intended for deaf aids.

The 'law' of these devices may be linear, log, semi-log, linear tapered, inverse log, inverse semi-log, or inverse linear tapered. The significance of these terms is explained in *Figure 20.4*. The main application of log law

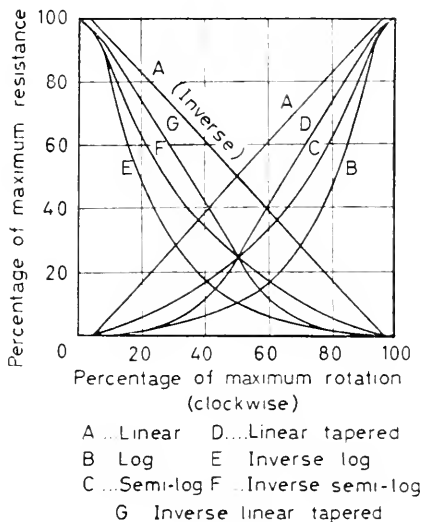


Figure 20.4

potentiometers is the volume controls of radio sets. The object here is to arrange that the subjective loudness obtained is proportional to the rotation of the control. In scientific work these components go some of the way towards providing a control roughly linear in dB's. Log law potentiometers may be made in the moulded track type by progressively varying the carbon content of the mix along the track length, and in the sprayed type by spraying the track a number of times, each time through a different specially shaped mask so that a non-uniform conducting layer is built up.

Wirewound variable resistors

With wirewound variable resistors of the 'radio' type under discussion, the law is nearly always linear though other laws may be had to special order. The range of values obtainable is typically 1 ohm–100 k, by way of the 1–2.5–5–10 progression, as for carbon components. The tolerances on nominal value are about 10 per cent. Manufacturing procedure is to wind nickel chrome wire on to a flat strip of insulating material, afterwards bending the strip round to the horseshoe shape necessary to fit into the insulating body of the component (*Plate 20.9*). The oxide film on the wire surface is relied upon for inter-turn insulation. Size for size, wirewound resistors have a much higher power rating than carbon—the popular size can dissipate 5

RESISTORS

watts—and this is probably their most important advantage. A possible disadvantage of the wirewound type is the fact that control takes place in discrete steps, whereas with a carbon component in good condition the resistance is varied smoothly.

The voltage rating of variable resistors, whether of wirewound or carbon type, ranges from about 250 V for the miniature types to 500 for the standard sizes.

Towards the end of their life, variable resistors tend to become 'noisy'. This is due to erratic contact between the slider and the track, the result of wear, or it may—in the case of high value wirewound components—be due to an intermittent open circuit in the track itself; the fine wire fails somewhere, perhaps due to corrosion, and the butting ends are now forced together, now parted, by slight movements imparted from the slider.

CAPACITORS

FIXED CAPACITORS

Fixed capacitors are available in values ranging from about 1 pF ($1 \text{ pF} = 10^{-12} \text{ F}$) to $2,000 \mu\text{F}$ ($1 \mu\text{F} = 10^{-6} \text{ F}$). In choosing a component one has to bear in mind not only the capacitance required but also the maximum potential difference that will occur between the plates of the capacitor and the nature of the dielectric material.

The dielectric material is determined to a considerable extent by the capacitance required. *Figure 21.1* shows approximately the capacitance

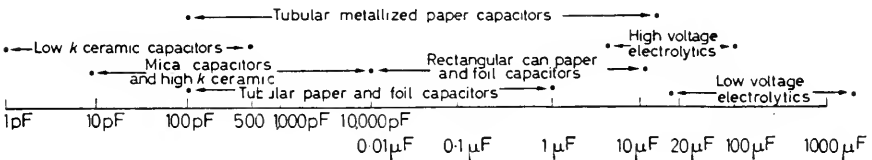


Figure 21.1

range occupied by capacitors employing the various common dielectrics, from which it is clear that for most values there are alternatives. To enable the reader to make an appropriate choice we shall mention briefly the characteristics of the various materials, but before doing so it is necessary to say a little about capacitors in general.

The current through a real capacitor when connected to a source of sinusoidally alternating voltage leads that voltage by an angle somewhat less than 90° , which implies that in addition to the true capacitive current there is a small component of current in phase with the applied voltage which represents energy delivered to the capacitor and not recovered, i.e. a 'loss'. The angle between the actual current vector and the ideal current vector is a measure of the loss occurring in the capacitor and is called δ (*Figure 21.2*).

A.c. bridges for the measurement of capacitors are commonly provided with two dials, both of which have to be set to the correct position in order to achieve balance; one is calibrated in capacitance, the other in ' $\tan \delta$ '. To assist readers who possess a capacitance bridge to check the condition of suspected capacitors of various kinds (and to show which kinds of capacitor are intrinsically 'lossy') very *approximate* values of $\tan \delta$ are given in this chapter.

Each dielectric material has a band of frequencies over which $\tan \delta$ is reasonably constant, but at the ends of the band the loss rises steeply. Electrobiology being a low-frequency technique, we are not concerned with losses at the upper end of these bands, all of which are much above the

CAPACITORS

electrophysiological spectrum: the low-frequency ends are, however, of considerable importance. The principal reason for the rise in loss at low frequencies with a particular dielectric is the increasing importance of the

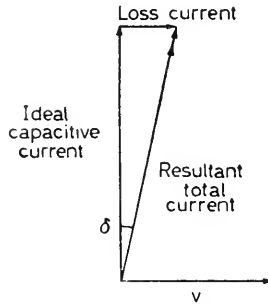


Figure 21.2

leakage current in comparison with the capacitive current. It follows that the leakage current of capacitors for *low-frequency* use gives a fair indication of the condition of the capacitor: this is a useful fact for the worker not possessing a capacitance bridge. Capacitor leakage resistance may be measured in a number of ways, of which the simplest is probably to charge the component and observe the rate at which it discharges through its own leakage with an electrostatic voltmeter (or high resistance moving-coil voltmeter in the case of low voltage electrolytics). The voltage will fall to $1/e$ of its original value in RC seconds, where C is in μF and R in megohms. As a guide to the values of leakage which are to be expected, approximate figures are included where applicable in this chapter. The leakage tends to be proportional to the capacitance, so the figures are in megohm- μF . To get the expected leakage resistance for a particular capacitor, divide the figure given by the capacitor value in μF .

Impregnated paper capacitors

In these capacitors the dielectric is paper impregnated with mineral oil, paraffin wax or petroleum jelly. The electrodes may be either of aluminium foil or may be made by spraying or vacuum-depositing a metal film directly on to the dielectric paper itself. The method of construction is to roll up a sandwich of foils and paper, or two layers of metallized paper into a cylindrical bundle, as indicated in the end views (*Figure 21.3*). The bundle is then enclosed in a tubular container or flattened and fitted into a rectangular one (*Plates 21.1* and *21.2*). It is clear that one or other electrode must be outermost with this construction and will form an electrostatic screen for the other. The connection to the outer electrode is marked by a black band on the container (*Figure 21.4*). The point is of importance in decoupling applications, where it is usual to earth the outside foil.

Maximum and working voltages for paper and foil capacitors range from hundreds to thousands, the most common ratings being 350 and 500. Metallized foil types range from 75 to 600. Tolerances are about ± 20 per cent. Values proceed by the 1-2-5-10 sequence, sometimes including 2.5 and 3.

FIXED CAPACITORS

Metallized foil capacitors are characterized by smaller size than foil and paper types for the same capacitance and working voltage, but also by greater leakage and loss. Reasonable figures are:

Foil and paper	$\tan \delta \approx 0.01$	Leakage resistance $\approx 5000 \text{ M}\Omega \cdot \mu\text{F}$
Metallized paper	$\tan \delta \approx 0.02$	Leakage resistance $\approx 100 \text{ M}\Omega \cdot \mu\text{F}$

The low leakage resistance of metallized paper capacitors may render them unsuitable for intervalve coupling purposes, and their recommended function

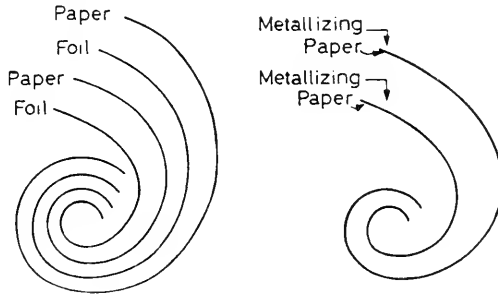


Figure 21.3

is decoupling and smoothing. The author has tried them for intervalve coupling and found that, in addition, they seem to generate noise, which makes them unsuitable in low-level stages. A useful attribute of metallized paper capacitors is a measure of self-protection against breakdown from over-

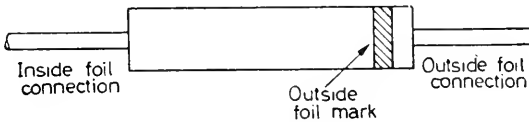


Figure 21.4

voltage; the fault current produces enough heat to evaporate the metallization in the damaged region, thus effectively removing that part of the component from circuit.

G. W. A. Dummer (*Fixed Capacitors*, London; Pitman) has drawn attention to an additional limitation with capacitors employing paper dielectric, which is that, however high the direct voltage rating of the component, any alternating voltage applied across it must never exceed about 350 peak. The reason is bound up with eventual burning of the paper as a result of repeated ionization of residual air bubbles within the component at each cycle of the applied voltage.

The temperature coefficient of paper capacitors is about $+0.02$ per cent per degree C.

Mica capacitors

Here the dielectric is mica and once again the electrodes may be either actual metal foils or may be metal films—commonly silver—applied direct to

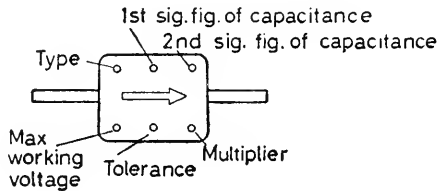
CAPACITORS

the mica by reducing an oxide powder of silver at high temperature (*Plate 21.3*). All mica capacitors have extremely low loss ($\tan \delta = 0.0003$) and the silver mica types have the added advantage that the process lends itself to an extremely close manufacturing tolerance, ± 2 per cent. In this way components whose actual values lie close to the nominal may be bought cheaply, because further sorting and selection is not required.

Mica capacitors are only made in low capacitance values (10–10,000 pF), otherwise their cost would be prohibitive. This means that they cannot easily be checked by examining the rate of self-discharge, since the leakage resistance (about 10,000 M Ω μ F) of the component is likely to be so high as to become comparable with that of the measuring device. However, mica capacitors are extremely reliable, and if they have not failed as a result of catastrophic damage such as would show up on a simple ohmmeter test, they are likely to be in good order. The temperature coefficient of foil and mica capacitors is better than ± 0.05 per cent per degree C; that of silver mica components better than ± 0.01 per cent per degree C.

Some mica capacitors are colour-coded, using up to 6 dots. The system is as follows:

British Mica capacitor colour code



Colour	Type	1st and 2nd significant figures of capacitance in pF	Multiplier	Tolerance	Volts
Black	—	0	1		
Brown	—	1	10		
Red	Silver mica	2	100	$\pm 2\%$	350
Orange	—	3	1,000		
Yellow	—	4	—		
Green	Foil and mica	5	—		750
Blue	—	6	—		
Violet	—	7	—		
Grey	—	8	—		
White	—	9	—		2,000
Gold	—	—	—	$\pm 5\%$	
Silver	—	—	—	$\pm 10\%$	
Black or Blank	—	—	—	$\pm 20\%$	

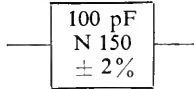
Low k ceramic capacitors (*Plate 21.4, left*)*

These are made over approximately the same capacitance range as mica capacitors (1–500 pF), and have a similar $\tan \delta$. Further, like the silver mica

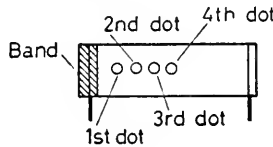
* k is the *permittivity*, or specific inductive capacitance, of the dielectric.

FIXED CAPACITORS

capacitor they can be made to extremely close tolerance, ± 1 per cent. Their most important characteristic, however, is that their temperature coefficients are accurately known. This means that in the design of RC filters and oscillators the temperature coefficients of the resistors may be allowed for by choice of suitable ceramic capacitors to go with them. The temperature coefficient range obtainable is $+0.01$ per cent per degree C to -0.075 per cent per degree C, and is stated on the component. The information may be printed on the component in some such manner as this



which means that the capacitance lies between 98 and 102 pF and the temperature coefficient is negative, 150 parts per million per degree centigrade. Alternatively, the component may be colour coded using the band-and-4 dot system. The colours have this significance:



Colour	Band. temp. coefficient parts/million/°C	1st and 2nd dots capacitance in pF	3rd dot, multiplier	4th dot	
				10pF or under	Over 10pF
Black	0	0	1	$\pm 2\text{pF}$	$\pm 20\%$
Brown	-30	1	10		$\pm 1\%$
Red	-80	2	100		$\pm 2\%$
Orange	-150	3	1,000		$\pm 25\%$
Yellow	-220	4	10,000		
Green	-330	5		$\pm \frac{1}{2}\text{pF}$	$\pm 5\%$
Blue	-470	6			
Violet	-750	7			
Grey	+30	8	0.01	$\pm \frac{1}{4}\text{pF}$ $\pm 1\text{pF}$	$\pm 10\%$
White	+100	9	0.1		

High k ceramic capacitors (Plate 21.4, right)

These are characterized by small physical size for their capacitance and voltage rating, which makes them useful for miniature equipment. They should be restricted to non-critical applications such as smoothing and decoupling, since their capacitance values are subject to serious fluctuations. Manufacturing tolerances are commonly of the order of $-20 +100$ per cent, and they have a very large negative temperature coefficient, about ten times bigger than that for any other kind of capacitor except the aluminium electrolytic. In addition they are inclined to be 'lossy' ($\tan \delta = 0.02$) and this, coupled with their small size, means that they are liable to heat up unless the alternating voltage across them is kept small. They are also rather leaky, of the order of $100\text{ M}\Omega \mu\text{F}$.

Electrolytic capacitors

Electrolytic capacitors are also characterized by large capacitance in a small volume. Though they are made in working voltages ranging from 6 to 600, in practice they fall into two important groups:

(1) Low voltage types, of working voltage less than 50, for cathode bias resistor bypass, and mains-derived d.c. heater supply smoothing. Capacitance values 20–2,000 μF .

(2) High voltage types for HT supply smoothing and decoupling, rated between 300 and 600 V, with capacitance values of 4, 8, 16, 32 or 50 μF .

In electrolytic capacitors the positive electrode is aluminium foil and the negative electrode is a paste of glycol and ammonium tetraborate to which connection is made by a second aluminium foil. The dielectric is a very thin layer of aluminium oxide covering the positive foil. This layer is formed by electro-deposition before the capacitor is assembled, and the component must always be wired up the correct way round so that the oxide film is not removed again, that is, the electrolytic capacitor is restricted to applications where the applied voltage is direct, or where, if there is a superposed alternating component, the net voltage never reverses.

The oxide film is extremely thin, of the order of 0.1μ , and about $1/100$ of the thickness of the paper used in impregnated paper capacitors. The capacitance is therefore very high for the size of the component, but so also is the leakage. The capacitance may be further increased by etching with acid the surface of the positive aluminium foil, thus increasing the surface area. This process does not give a commensurate increase in leakage. The leakage resistance of plain foil electrolytic capacitors is about $6 \text{ M}\Omega\text{-}\mu\text{F}$, and of the etched foil type, $20 \text{ M}\Omega\text{-}\mu\text{F}$.

Electrolytic capacitors are somewhat imprecise components; manufacturing tolerances are commonly -20 , $+50$ per cent, and the temperature coefficient is large and positive, about 0.2 per cent per degree C. They are also extremely liable to loss— $\tan \delta = 0.15$ —and this sets a limit to the magnitude of the alternating component of voltage which may be allowed across them. If this is excessive, the heat generated within the component will volatilize the electrolyte paste and generate a pressure which may cause the component to explode. When this happens the noise is extremely startling and the remainder of the interior of the apparatus becomes plastered with fragments of electrolytic capacitor.

In most electrolytic capacitor applications the a.c. component is not serious, but there is one important exception; this is in capacitor-input power-rectifier circuits. With some capacitors the maximum permissible a.c. is marked, e.g. '130 mA. a.c. Max', but where this is not done the manufacturers should be consulted. In the highest class of design a paper and foil capacitor is used in this position. Electrolytic capacitors are frequently assembled two or three to a single container; when this is done there is often a printed notice to say that one or other section is intended for high-a.c.-ripple-duty.

In addition to the maximum direct working voltage rating, electrolytic capacitors are also often marked with a rather larger 'surge' rating. This is a voltage which may be applied to the component for a short time without damaging it. The point is important because in many pieces of electronic

VARIABLE CAPACITORS

apparatus using non-thermionic rectifiers, or thermionic rectifiers of the directly heated kind, the HT is available before the valves have warmed up to consume it, and during this transient period the HT voltage 'surges' to an excessive value.

The mechanical form of electrolytic capacitors is either tubular, with a wire emerging at each end (see *Plate 21.5*) or enclosed in a cylindrical can, secured to the apparatus by a single large nut or by a clip, and having all the connections at one end. Where several capacitors are enclosed in one case the latter construction is usual, and the capacitors often have a common negative pole which is internally connected to the can itself.

Future developments

At the time of writing there are two further types of capacitor of great promise which are not widely used, probably because at present they are too expensive.

Polystyrene film capacitors—These have low loss ($\tan \delta = 0.0005$) and phenomenally low leakage, about 30 times less than a mica capacitor and 100 times less than impregnated paper and foil. They are made from 100 pF to 4 μF and should be ideal for coupling purposes in low-level stages.

Tantalum electrolytic capacitors—These are even smaller for their size and working voltage than aluminium electrolytic components. One 50 μF 70 V unit is about the size of a stack of half-a-dozen halfpennies. In addition they have better $\tan \delta$, about 0.05, and extremely low leakage, about a hundred times better than conventional electrolytics.

VARIABLE CAPACITORS

Variable capacitors are of two types: (1) the truly variable type intended for frequent readjustment, employing air as dielectric, and familiar to anyone who has looked inside a wireless set as the 'tuning condenser'; (2) 'preset', screwdriver-adjusted types, otherwise 'trimmers', meant to be initially set up to a correct value and left there.

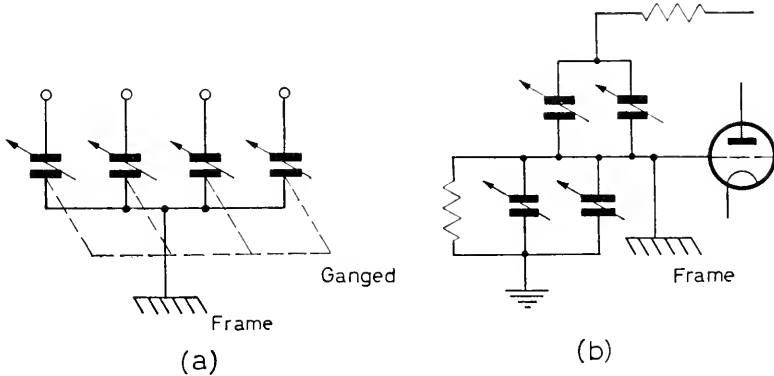


Figure 21.5

True variable capacitors

These are useful for the accurate control of variable-frequency filters and oscillators. The most common size is variable up to 500 pF and may be had

CAPACITORS

in gangs up to 4. Thus by parallel connecting in pairs one can make $2 \times 0.001 \mu\text{F}$ max. for controlling Wien bridge type oscillators and filters, but a word of warning is necessary—the frame of these capacitors is common to one set of plates in each gang (*Figure 21.5*). Thus when they are used in Wien bridges the frame is live and the component must be mounted on insulating pillars and operated via an insulating shaft. This sharing of one terminal means ganged variable capacitors cannot be used for tuning parallel T 's. It is necessary to devise some other method of mechanically coupling individual variable capacitors, such as the use of nylon gears.

Trimmers

These commonly have maximum capacitances of about 50 pF; a group are shown in *Plate 21.6*: (a) is virtually a tiny air dielectric variable capacitor; (b) is another air-dielectric type, range 3–30 pF, obtained by screwing in and out of one another two sets of concentric annular plates; (c) is a low k ceramic type, range 5 to 40 pF, with a stated temperature coefficient (either -750 p.p.m./ $^{\circ}\text{C}$ or $+100$ p.p.m./ $^{\circ}\text{C}$); (d) is a 'mica compression' type. It is in effect a foil and mica capacitor with arrangements for varying the pressure which holds the assembly together and thus adjusting the capacitance value. The pressure is controlled by turning the screw.

CHOKES AND TRANSFORMERS

These are the practical versions of the theoretical self and mutual inductances*. It has been intimated before, in Part I, that it is a good plan to keep inductors of both kinds out of electronic equipment as far as possible, and that one reason for this is that whereas one can go into a radio shop and buy a resistor or capacitor with for practical purposes any value over a range of several orders, the same is not true of inductors. One may get what one wants by improvization; a particular radio tuning coil or old intervalve transformer may happen to have a required self inductance, but if this is so it is fortuitous. In general, if one wants inductors one has to make them oneself.

There are some important exceptions: the mains transformers and chokes sold as replacements for radio sets are quite suitable for the power packs of electronic equipment, and it is often feasible to press the multi-ratio types of 'output transformer'—for coupling power valves to loudspeakers—into services for which they were never intended.

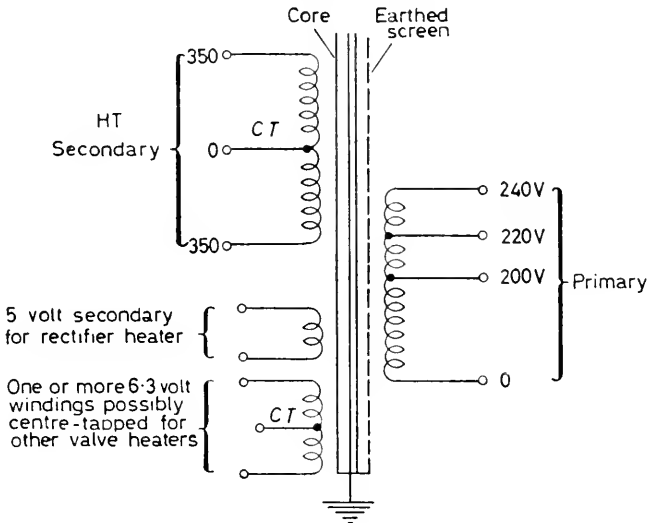


Figure 22.1

MAINS TRANSFORMERS

The most important class of these (Figure 22.1) have about 4 distinct windings:

(1) A primary winding, for connection to the mains and provided with taps at, say, 200, 220 and 240 V for adjustment to the appropriate local voltage.

* Except in the world of LC filters, where the inductances retain their academic name—inductors.

CHOKES AND TRANSFORMERS

(2) A high voltage secondary winding for the generation of HT. This may be either for use with half-wave or bridge rectification, in which case there is a single winding with two ends brought out, or for full-wave rectification, in which case the winding has twice as many turns and both ends and the centre-tap are brought out. Full-wave HT secondaries are more common. A half-wave or bridge winding which delivers 350 V R.M.S. at a max permissible load of 120 mA would be described as 350 V, 120 mA. A full-wave counterpart would be designated 350-0-350 V, 120 mA. HT secondaries range from about 150-0-150 to 500-0-500 for voltage and from about 60 mA to 350 mA. The most popular and generally available sizes are probably 350-0-350 V 60 mA and 350-0-350 V 120 mA. It should be remembered that the rectified HT voltage of a 350-0-350 V secondary, using the usual capacitor input scheme, and on light load, is much higher than 350. First the R.M.S. output on light load of a winding designed to give 350 on full load is likely to be about 380. Secondly the rectifier system will deliver the peak voltage, $(2)^{1/2} \times 380 = 540$ V. The point should be borne in mind when choosing the voltage ratings of capacitors.

(3) A low voltage secondary winding delivering 5 V at about 2 amps. This is intended for the heater of a thermionic rectifier where one is used (see Chapter 24).

(4) One or more 6.3 V windings delivering between 1 and 3 amps maximum for valve heaters generally and for such accessories as dial lamps. These windings may be provided with centre-taps; these should be earthed, the purpose being to balance out the electrostatic interference caused by the two leads comprising the heater wiring.

In addition there may be an odd wire—frequently uninsulated—and this goes to an electrostatic screen between the primary and secondary windings. It should be earthed. The idea is to keep mains-borne interference at radio frequency (e.g. diathermy apparatus) out of the equipment.

SPECIAL MAINS TRANSFORMERS

The autotransformer

This is shown in *Figure 22.2*, from which it is clear that the component

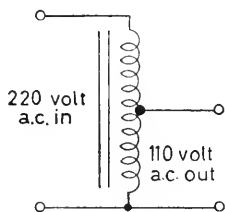


Figure 22.2

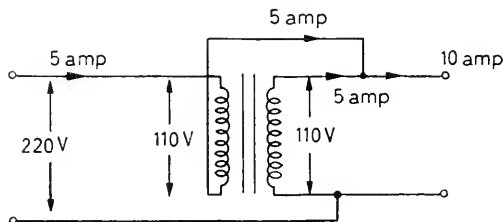


Figure 22.3

possesses only one winding. Autotransformers are commonly used for running American 110 V equipment from British 220-odd V mains, or vice versa. Not only is the autotransformer advantageous in that—having only one winding—it is easier to make, but also the component can be smaller. This

is easily seen by re-drawing the device as in *Figure 22.3*. Suppose the load current is 10 amps, the power which the transformer has to handle if it is of the conventional 'double wound' type is $110 \times 10 = 1,100$ watts, neglecting losses. By regarding the autotransformer as a double wound component as in *Figure 22.3*, it is easily seen that the power handled is only 550 watts. In general, the nearer the voltage transformation ratio is to unity the greater is the advantage offered by autotransformation.

The variac

This (*Figure 22.4*) is an autotransformer of continuously variable turns-ratio for interposition between a piece of electronic apparatus and the mains: with its help, fluctuations in mains voltage may be manually corrected.

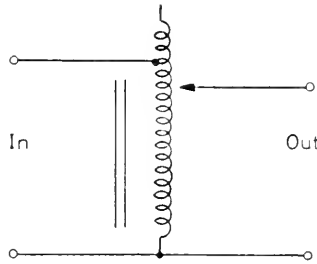


Figure 22.4

Another important use is for subjecting apparatus possessing automatic voltage regulating devices to dummy mains voltage changes, to check the operation of the regulator. The output voltage is variable from zero to some 20 per cent above the input, and variacs can be had in a range of sizes enabling them to handle maximum powers between 170 and 25,000 watts.

EHT transformers

Similar to mains transformers, but the HT winding is invariably intended for half-wave rectification and delivers several thousand volts at a few milliamperes; one end is often internally earthed. Any LT windings present have usually to be insulated to withstand approximately twice the peak EHT voltage.

Constant-voltage transformer

These devices are capable of maintaining an output R.M.S. voltage constant within 1 per cent, despite ± 15 per cent changes in input voltage. They have two secondary windings connected in series opposing, a main winding excited by a core which is continuous and a subsidiary winding excited by a core possessing an air gap. Matters are arranged so that over the working range of input voltages the gapless core runs into saturation, whereas the gapped one of course does not. The outputs have the form of *Figure 22.5*, and sum to produce a rather constant output, as shown. Disadvantages of the constant-voltage transformers are that, unless special filters are provided

CHOKES AND TRANSFORMERS

the output waveform is distinctly non-sinusoidal, also they are rather large for the power output they produce.

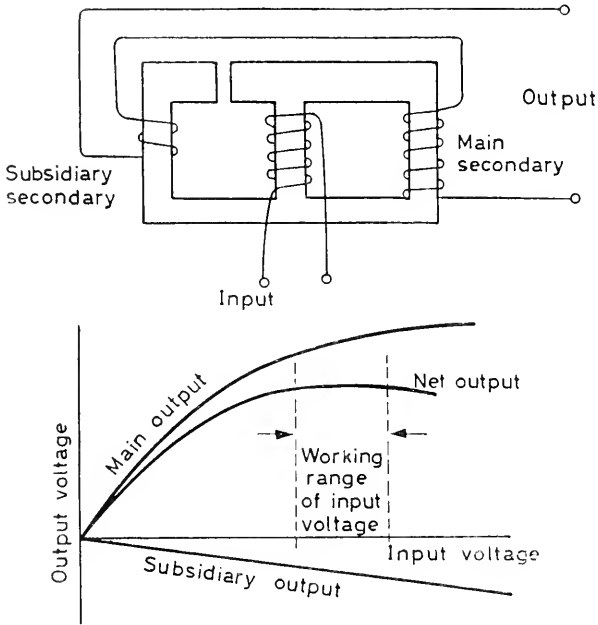


Figure 22.5

SMOOTHING CHOKES

These range in inductance from 5 h to 75 h, and in current from about 25 mA max to 500 mA max, but it is necessary to point out that the low current ratings tend to go with the high inductances and vice versa. A 75 h 500 mA choke would be extremely large. The most usual values are 10 h and 20 h at about 120 mA. In describing swinging chokes the inductance without polarizing d.c., and with full d.c., are both quoted.

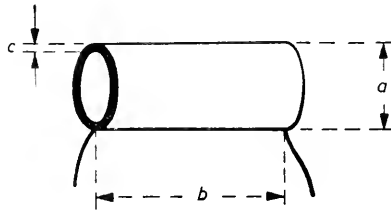


Figure 22.6

DESIGN OF SMALL AIR-CORED CHOKES

Small inductances up to a few tens of millihenries may be made by winding a simple coil of wire on some kind of bobbin of insulating and non-magnetic material (Figure 22.6). The winding is then cylindrical and if a is the average

diameter, b the length and c the thickness of winding, all dimensions in inches, then the self-inductance in microhenries is approximately given by¹

$$L = \frac{0.2a^2n^2}{3a + 9b + 10c}$$

DESIGN OF SMALL IRON-CORED CHOKES AND TRANSFORMERS

This is a highly elaborate subject, and a specialist field in itself. The ability to design efficient small transformers for power, signal and pulse applications, comes to most people only after many years of experience; readers who feel drawn to the subject may care to look at an excellent book by Macfadyen². A complete account of the subject cannot possibly be included here; nevertheless the author proposes to outline a procedure, trial-and-error rather than synthetic, by which the reader should be able to make himself small transformers and chokes which, though perhaps not the best that can be done, should prove sufficient for purposes such as LC oscillators and filters.

The method is to use an interesting magnetic material 'Ferroxcube', made by the Mullard Company. Unlike most transformer cores, which are metals and are alloys of nickel and iron, this material is a crystalline substance of the form MFe_2O_4 where M is a divalent metal. Here we are interested in a grade of Ferroxcube called A4 having the formula $MnFe_2O_4$. The advantages of this material are:

(1) Extremely low losses permit use up to frequencies of at least half a megacycle, much higher than are feasible with the nickel irons. Ferroxcube is less efficient than the latter at very low frequencies because the saturation flux density is lower; nevertheless, one can reckon to use this single material successfully at any frequency of electrobiological interest, so one is relieved of the need for selecting one's magnetic medium.

(2) The manufacturers make a range of 11 moulded E and 3 I members in Ferroxcube A4 with which it is possible to make up 14 different sizes of core, either by using $2E$'s or an E and an I ; one of these is likely to suit one's requirements. They also can supply winding bobbins to go with them. Assembly is therefore merely a question of winding the bobbin, slipping the core round it and cementing the whole together. This saves the labour of making one's own bobbin and assembling a core from laminations, as is usually necessary when using nickel iron cores. (*Plate 22.1.*)

(3) The information provided by Mullard Ltd on these cores and the fact that 'iron losses' are usually negligible make designs very straightforward.

We have to distinguish two cases, according to whether or not the transformer or choke cores operate with d.c. polarization; we consider the unpolarized case first.

Ferroxcube transformers and chokes, without d.c. polarization—Chokes (Inductors)—It is required to produce a choke of inductance L which will have a resistance not greater than R_{\max} and at a minimum frequency F_{\min} will operate satisfactorily with terminal alternating R.M.S. voltages up to \bar{V}_{\max} .

First, choose a core: at first one simply has to make a guess here. It is not likely that this is the core one will finally use, but no matter—with

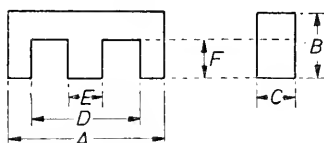
CHOKES AND TRANSFORMERS

TABLE 1*

Dimensions of Ferroxcube *E* Cores (in mm). These are Subject to Minor Variations

<i>Type</i>	<i>A</i>	<i>B</i>	<i>C</i>	<i>D</i>	<i>E</i>	<i>F</i>	<i>Remarks</i>
FX 1052	12.7	6.6	3.18	9.53	3.3	4.95	(Centre leg circular in cross-section)
FX 1238	25.4	9.53	6.47	19	6.35	6.35	
FX 1105	30.0	12.5	12.0	20.25	9.75	7.5	
FX 1239	34	13.1	7.87	24.6	11.1	8.5	
FX 1073	35.1	9.9	12.0	23.9	9.3	5	
FX 1007	42.3	22.5	8.9	28	11.7	16.2	
FX 1608	42.3	22.5	11.7	28	11.7	16.2	
FX 1186	42.5	18.5	9.55	27.5	9.6	11.0	
FX 1314	54.5	30.0	12.0	34.5	12.0	22.0	
FX 1384	55.4	29.4	12.7	38.1	12.7	20.4	
FX 1315	41.7	23.5	15.0	27.0	12.0	17.5	(Centre leg circular in cross-section)

E cores



I cores



* By Courtesy of Messrs Mullard Ltd.

Dimensions of Ferroxcube *I* Cores (in mm). Subject to Minor Variations

<i>Type</i>	<i>A'</i>	<i>B'</i>	<i>C'</i>
FX 1053	12.7	3.05	3.96
FX 1106	30.0	5.0	12.0
FX 1107	41.0	6.0	9.0

Data on Complete Cores, that is, 2 *E*'s or *E* and *I*

<i>Type numbers</i>	<i>l</i> (cm)	<i>A</i> (cm ²)	<i>v</i> (cm ³)	<i>Turns for one millihenry</i>	
FX 1052	FX 1053	2.1	0.1	0.21	62
FX 1052	FX 1052	3.1	0.1	0.31	72
FX 1238	FX 1238	4.8	0.4	1.9	30
FX 1105	FX 1106	4.1	1.2	4.9	18.5
FX 1105	FX 1105	5.6	1.2	6.7	19.5
FX 1239	FX 1239	6.3	0.78	5.0	25.4
FX 1073	FX 1073	5.0	0.68	5.0	22.8
FX 1007	FX 1107	6.8	1.05	7.1	22
FX 1007	FX 1007	10.0	1.05	10.5	27
FX 1608	FX 1608	10.0	1.4	14	24.2
FX 1186	FX 1186	8.7	0.9	10.6	23.7
FX 1314	FX 1314	13.8	1.4	23	30
FX 1384	FX 1384	13.4	1.6	27	26.5
FX 1315	FX 1315	10.5	1.13	17	22.5

DESIGN OF SMALL IRON-CORED CHOKES AND TRANSFORMERS

experience one's guesses improve. From *Table 1* find out the number of turns for one millihenry (this is our $(K_2 l)/(\mu K_3 K_1)$ in Chapter 4 all worked out for us), and thus compute the number of turns N required, remembering that $L \propto (\text{turns})^2$: also from *Table 1* find out the 'window area' of the proposed core, through which the turns will pass and, for safety, take two-thirds of it (this is to allow for imperfections in one's winding technique). From *Table 2* find out the thickest gauges of wire which will allow one to get the

TABLE 2*
Copper Wire Information

<i>S.w.g.</i>	<i>Resistance (ohms per yard)</i>	<i>Maximum turns per cm² of window area</i>	
14	0.0048	17	} Double cotton covered
16	0.0075	26	
18	0.013	44	
20	0.024	89	} Enamelled and double silk covered
22	0.039	137	
24	0.063	212	
26	0.094	343	} Enamelled and single silk covered
28	0.140	490	
30	0.20	674	
32	0.26	852	
34	0.36	1,135	
36	0.53	1,550	
38	0.85	2,250	
40	1.33	3,170	
42	1.91	3,910	
44	2.99	5,320	
46	5.31	7,650	

For wire insulated with enamel alone, the turns per cm² will be somewhat greater, particularly in the fine gauges.
* Derived from reference 1.

required number of turns into the space provided: by multiplying the number of turns by the mean length of turn, estimate the total length of wire required. From the resistance per yard for the gauge of wire in use, compute the resistance of the proposed winding; if this is less than R_{\max} , all is well so far, if it exceeds R_{\max} , try the whole thing again with a bigger core. Assuming the resistance criterion is satisfied, it remains next to investigate the amount of magnetization that will be produced. We had in Chapter 4 that the instantaneous e.m.f. across a pure inductance = $K_3 N (d\phi/dt)$. If this equals $V \cos \omega t$, then, integrating, $(V/\omega) \sin \omega t = K_3 N \cdot \phi$

$$\text{or} \quad \frac{V}{\omega_{\min}} = K_3 N \phi_{\max}$$

The maximum flux density, B_{\max} determines the degree of magnetization produced.

$$B_{\max} = \frac{\phi_{\max}}{\text{Cross-sectional area of magnetic circuit}} = \frac{V}{\omega_{\min} K_3 N A}$$

We are interested in the R.M.S. value of coil voltage, \bar{V} , so

$$B_{\max} = \frac{(2)^{1/2} \bar{V}}{\omega_{\min} K_3 N A} = \frac{\bar{V}}{(2)^{1/2} \pi K_3 N A f_{\min}}$$

CHOKES AND TRANSFORMERS

In the c.g.s. system of units, K_3 is equal to 10^{-8} , and B_{\max} comes out in gauss.

$$B_{\max} = \frac{10^8 \bar{V}}{4.45 N A F_{\min}} \text{ gauss}$$

where A is the cross-sectional area of the core in cm^2 . In a practical core A is the area of the centre limb. Work out B_{\max} for the core chosen: the answer ought to be substantially less than 2,000, otherwise the Ferroxcube will come too near saturation. If B_{\max} is greater than 2,000, try again with a bigger core.

If B_{\max} is much less than 2,000 and R comes out to much less than R_{\max} it may be worthwhile to try again with a smaller core. The object is usually to arrive at a design using the smallest and cheapest core which will satisfy requirements.

Transformers without d.c. polarization—It is required to produce a 1 to n transformer to work between a generator of internal resistance r and a load R_L , down to a turnover frequency F_{\min} . Maximum R.M.S. generator voltage, \bar{V}_{\max} . It is hoped that when the design is completed the winding resistances will be small compared with the generator and load resistances with which they are in series (Figure 22.7). If this is the case then the primary inductance required is roughly given by arranging that at F_{\min} the reactance of L_P equals the parallel resistance of r and R_L/n^2 , that is

$$\omega_{\min} L_P = \frac{r \cdot \frac{R_L}{n^2}}{r + \frac{R_L}{n^2}}$$

or

$$L_P = \frac{r \cdot \frac{R_L}{n^2}}{\left(r + \frac{R_L}{n^2}\right) 2\pi \cdot F_{\min}}$$

(cf. Chapter 4, expression for lower turn-over frequency of signal transformer.)

Work out L_P and choose a core. Find out the number of primary turns required to secure L_P and thus a suitable (i.e. thickest possible) gauge of primary wire, taking two-thirds of the window area as before and allowing half for the primary winding and half for the secondary. Estimate the resistance the primary winding is likely to have—as for chokes—and see that it is less than, say, 10 per cent of r . If not, try a bigger core. Multiply the primary turns by n to get the secondary turns, and thus compute a likely secondary wire gauge and hence secondary resistance. Check that the secondary resistance is less than about 10 per cent of the load resistance. If not, begin all over again with a bigger core. If all seems to be well, compute B_{\max} as for chokes (but bear in mind that the applied voltage is not \bar{V} but $\bar{V} [(R_L n^2)/(R_L n^2 + r)]$ (Figure 22.7) and check that the value is below 2,000. If B_{\max} is safely within bounds and the estimated winding resistances are much less than the external generator and load resistances with which they are in series, try re-designing with a smaller core.

Transformers and chokes with d.c. polarization—When the current through

DESIGN OF SMALL IRON-CORED CHOKES AND TRANSFORMERS

a choke or the primary winding of a transformer contains a direct component which is not balanced out, as happens in, e.g., a single-sided instead of a push-pull output stage, the iron experiences a mean magnetization, or bias, towards saturation and the permeability is found to be reduced: this causes a similar reduction in the self inductance. The remedy is to include an air

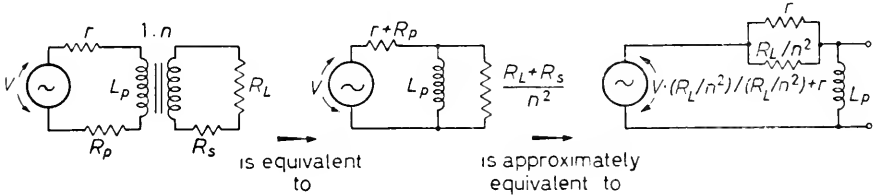


Figure 22.7

gap in the magnetic circuit, which increases the absolute reluctance of the magnetic circuit and hence reduces the polarizing flux which is the cause of the trouble. The consequent improvement in permeability reduces the incremental reluctance of the iron and, if the air gap is not too great, actually reduces the net incremental reluctance of the whole magnetic circuit. It is on the incremental reluctance that the inductance depends (Figure 22.8): the

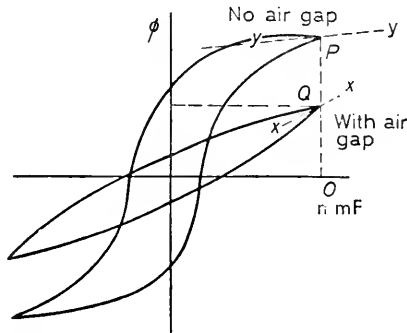


Figure 22.8. With no air gap a given mF produces a flux proportional to OP , representing a certain reluctance. The incremental reluctance is proportional to the reciprocal of the slope of yy , and is large. Introducing an air gap reduces the flux from OP to OQ , an increase in absolute reluctance; but the core material comes out of saturation, xx is steeper than yy , representing a fall in incremental reluctance

problem is to know how great the air gap ought to be. Fortunately the answer may be easily found from a design procedure due to Hanna³. A Hanna design curve for A4 Ferroxcube which is suitable for any value of B_{max} below 1,000 gauss is given in Graph 34. The method is to choose a core and to work out

$$\frac{LI^2}{v}$$

where L is the inductance required in henries, I is the polarizing current in milliamps and v is the volume of the core in cm^3 , then on the graph mark off on the curve a point opposite the value of LI^2/v just found. This point

CHOKES AND TRANSFORMERS

gives the required air gap, d , as a proportion of the total length of the magnetic circuit, l . By dropping down from this point to the horizontal axis one then reads off a value of N/l , where N is the number of turns required and l is the length of the magnetic circuit in cm. As l and l are known, N can be worked out. From the known minimum frequency and maximum primary voltage to be employed, the value of B_{\max} may be calculated. The Hanna design curve given is satisfactory provided B_{\max} does not come out to more than 1,000 gauss. If B_{\max} works out very different from 1,000 gauss, try another core.

Having found a core, gap and number of turns which gives the required inductance at the anticipated polarizing current, the rest of the design procedure—calculation of the winding resistances—is the same as for unpolarized chokes and transformers, except that in the case of transformers it is probably best to allot rather more window area to the primary than to the secondary winding.

REFERENCES

- ¹ *Amateur Radio Handbook* Amateur Radio Relay League
- ² MACFADYEN, R. A. *Small Transformers and Inductors* London; Chapman and Hall
- ³ HANNA, C. R. *J. Amer. Inst. elect. Engrs.* 46 (1927) 128

NON-THERMIONIC DIODES

Non-thermionic diodes rely for their action on the rectifying properties of junctions between certain pairs of materials. The important pairs are:

Copper-Copper oxide
Selenium-Iron
P Germanium-N Germanium
P Silicon-N Silicon

The letters P and N associated with germanium and silicon denote the presence of minute quantities of specially introduced contaminants which make rectification possible. In each case the direction of easy current flow is from the first to the second material. Terminals are labelled either by colours or signs, as shown in *Figure 23.1*.

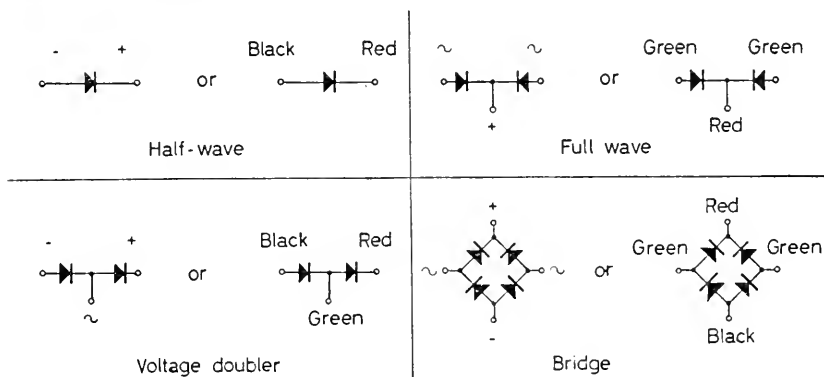


Figure 23.1

Each of these four junctions has associated with it, at a particular temperature, two important ratings:

(1) A maximum mean forward current, such that the heat produced shall not damage the junction. This current is roughly proportional to the junction area.

(2) A maximum backward voltage or 'Peak Inverse Voltage', such that the electric field produced shall not damage the junction. Diodes are often used in applications where the P.I.V. for a single junction would be exceeded. Hence a complete diode unit consists in general of a number of junctions in series. Consequently complete rectifying assemblies are a 'pile' or 'stack' structure whose length is an indication of the voltages, and whose thickness is a measure of the current with which they are intended to be used.

Diodes intended for signal rectification generally have only one junction and are usually referred to as 'diodes'. Diodes intended for a.c. power

rectification generally have the stack structure and are usually referred to as 'rectifiers'.

Unfortunately the voltage rating of a rectifier may be stated in either of two ways, and this sometimes leads to confusion. Consider the circuit in *Figure 23.2*, where we have a rectifying system of the half-wave capacitor-

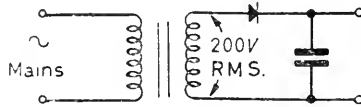


Figure 23.2

input type. On light load this arrangement gives an output voltage of approximately $(2)^{1/2} \times 200 = 283$. A little thought reveals that the peak inverse voltage across the rectifier occurs in the middle of the non-conducting part of the cycle and is 566 V. Sometimes it is explicitly stated that the P.I.V. of a rectifier is, say, 600 V, rendering it suitable for this application, but it may also be described as a '200 V' rectifier, referring to the maximum transformer R.M.S. voltage. In looking at manufacturers' literature it is necessary to be very careful which figure is being quoted.

POWER RECTIFIERS

These fall into three groups:

(1) Types for supplying HT to valves, a few hundreds of volts at a few tens of milliamps.

(2) Types for supplying 'extra HT' to cathode ray tubes, image-converter tubes, photomultiplier photocells and the like. These deliver a few thousands volts at a few milliamps.

(3) LT types, for making a few volts at a few amps. Important uses for these are accumulator charging, and the generation of direct heater current from the mains.

HT rectifiers

Some HT rectifiers are shown in *Plate 23.1*. Nos. 1, 2 and 3 are selenium-iron and are rated at 120, 60 and 30 mA '125 V', i.e. the P.I.V. is about 350 V. Notice the cooling fins; these components should be mounted in the attitude shown so that optimum convection occurs. The forward resistance of these units is such that about $\frac{1}{2}$ of the no-load output voltage is lost on full load. Nos. 4 and 5 are copper-copper oxide and are contact-cooled, that is, they are bolted against the chassis and the heat generated is removed by conduction. They are rated at 120 and 60 mA respectively, 125 V, so they correspond to the selenium iron units above them. No. 6 is a germanium junction type. *Provided it is kept properly cool (<25°C)* by suitable siting in the apparatus and by bolting it firmly down to a metal sheet, so that the heat generated is conducted away, this small unit can handle 1 A and withstand a P.I.V. of 200. Further, the makers claim a forward voltage drop less than 10 per cent of that caused by other types of rectifier. Nos. 7, 8 and 9 are silicon junction types of even more astonishing performance. 7 and 8



Plate 21.2 Group of metallized foil capacitors. These are 10 μ F, 150 V, 0.1 μ F, 350 V, and 0.001 μ F, 400 V

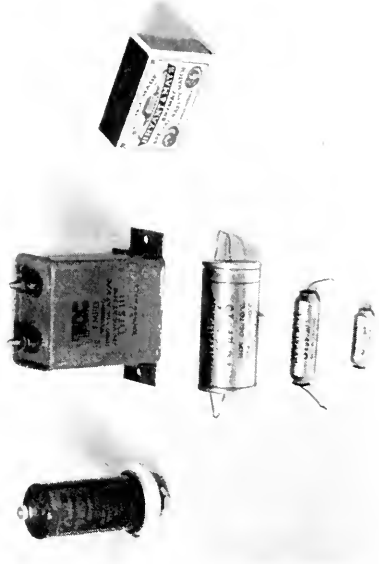


Plate 21.1 Group of foil-and-paper capacitors. The tubular capacitor standing upright is for EHT smoothing, and is 0.0005 μ F, 17 kV. The others are, reading downward, 1 μ F, 1,000 V, 1 μ F, 350 V, 0.1 μ F, 350 V, 0.01 μ F, 500 V

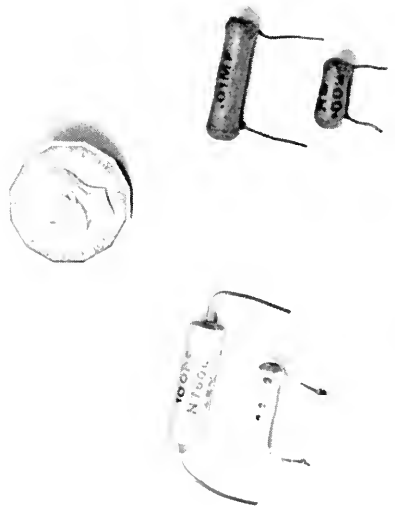


Plate 21.4 Ceramic capacitors. On the left, low k. On the right, high k

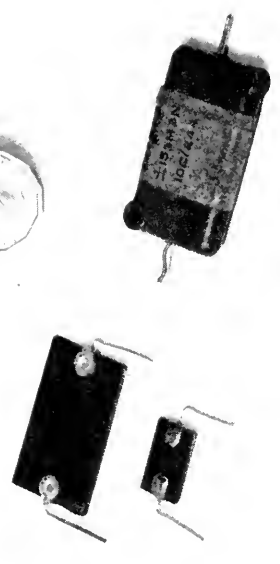


Plate 21.3 Mica capacitors. The pair on the left are silver mica. The component on the right is mica and foil

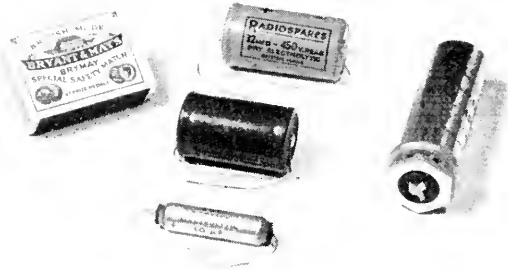


Plate 21.5 Group of electrolytic capacitors. On the right, a 'canned' capacitor, 32 μ F, 450 V; centre, beginning at the top, 32 μ F, 450 V, 16 μ F, 450 V, 50 μ F, 12 V

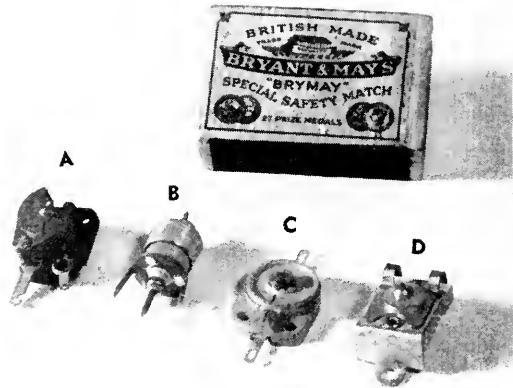


Plate 21.6 Group of trimmers

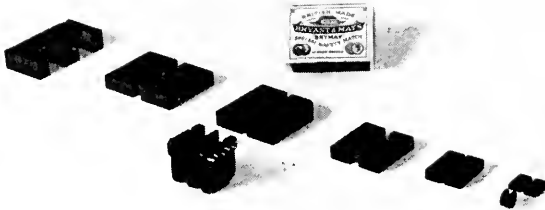


Plate 22.1 Group of Ferrocube E cores and bobbins

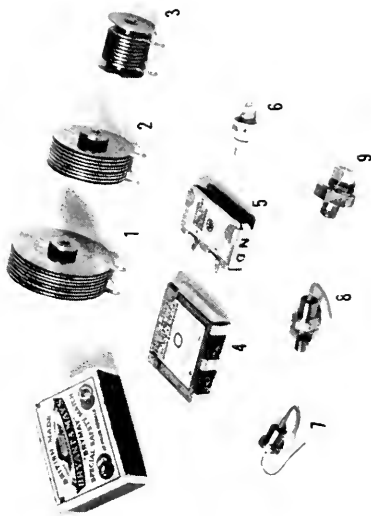


Plate 23.1 Non-thermionic HT rectifiers



Plate 23.2 Non-thermionic EHT rectifiers



Plate 23.3 Bridge-connected LT rectifier



Plate 23.4 Non-thermionic signal diodes



Plate 25.2 International octal, miniature octal and miniature-heptal valveholders

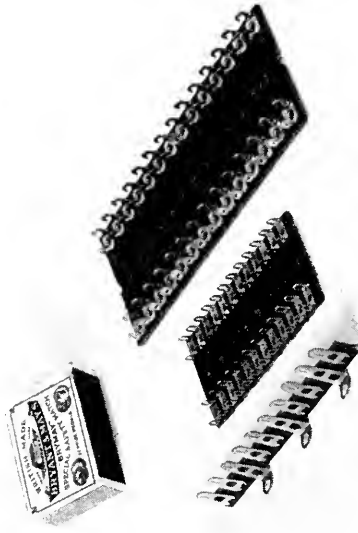


Plate 25.4 Tagboards and tagstrip

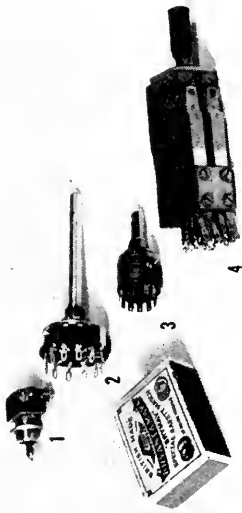


Plate 25.1 Group of switches



Plate 25.3 Fine control, pointer, and 'compromise' knobs

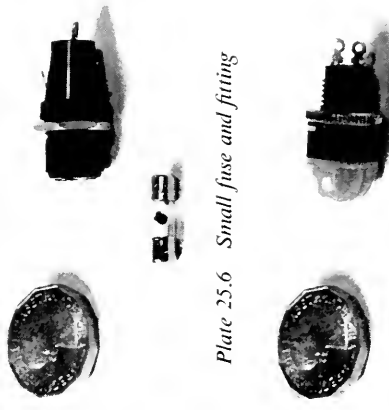


Plate 25.6 Small fuse and fitting



Plate 25.7 Indicator lamp and fitting

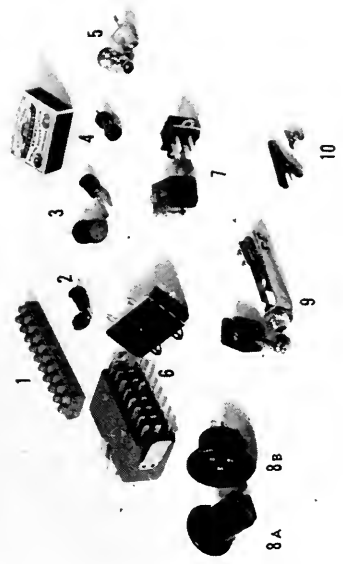


Plate 25.5 Selection of connectors

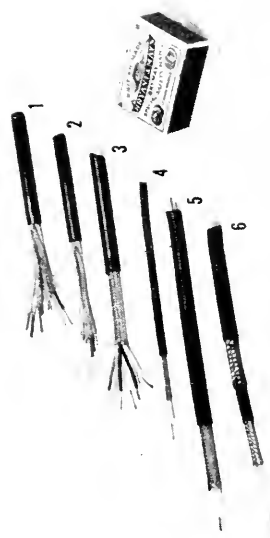


Plate 25.9 Group of cables

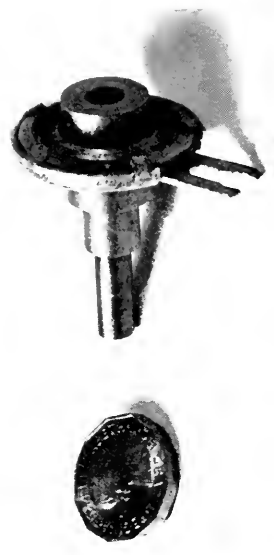
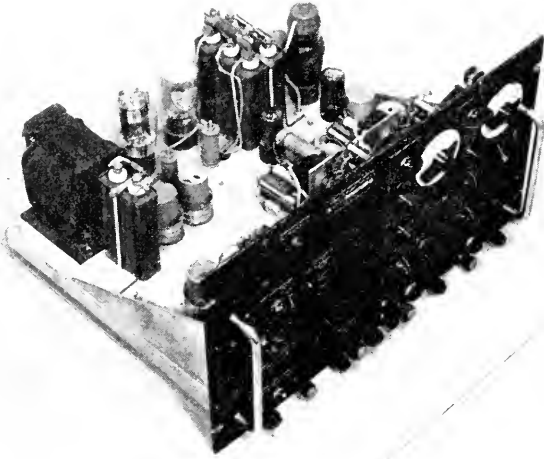


Plate 25.8 Epicyclic slow-motion drive



6.54

Plate 26.1 Chassis and panel construction. The panel width is 19 in.



Plate 26.2 19 in. rack-mounted equipment



Plate 26.3 14 in. chassis and panel construction. Three blank units, each comprising chassis, panel, and dust cover are shown mounted on a rack

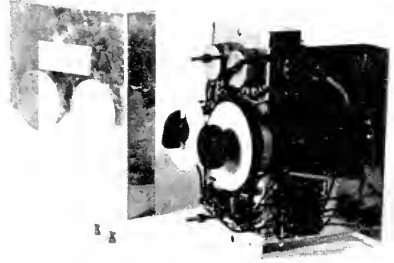
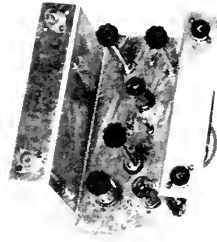


Plate 26.4 The author's 'inside-out' construction. Upper, back view; lower, front view

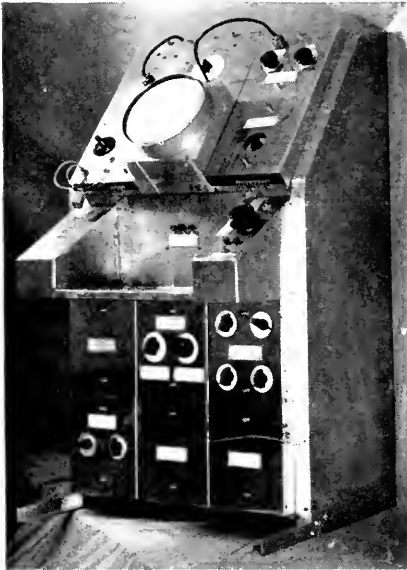


Plate 26.5 Prototype electrophysiological apparatus embodying 'inside-out' construction

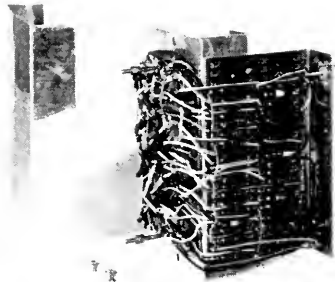
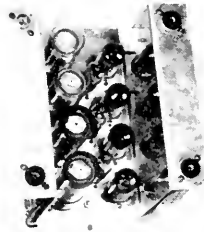


Plate 26.6 Another example of 'inside-out' construction



Plate 27.1 Foot-operated 3 ft. guillotine



Plate 27.2 Swing-beam folding machine

SIGNAL DIODES

have P.I.V.s of 200 V and can pass respectively 2 A and 8 A if the temperature is kept down to 25°C, and—an important point with silicon—will operate in a temperature of 100°C, when they can pass 1 and 4 A respectively. No. 9, by a different manufacturer, only passes a maximum of $\frac{1}{2}$ A in an ambient temperature of 25°C, falling to 150 mA at 100°C; but a P.I.V. of 400 V is possible.

EHT rectifiers (Plate 23.2)

The smaller object is a rectifying stack of the selenium-iron type rated at 1 mA, 600 V. The larger is copper-copper oxide, mean forward current 8 mA, 760 V. Voltage rating for voltage rating, selenium rectification is seen to produce a considerably shorter component.

LT rectifiers

Nowadays these seem usually to employ selenium rectification. LT rectifiers are made up by stacking junctions or 'plates' on to a threaded rod, securing them with nuts and using the ends of the rod as a fixing bolt. The object of the stacking is to secure appropriate voltage rating and rectifier function. Units are available ready connected up in any of the configurations of *Figure 22.1*, of which the bridge connection is the most important. The reverse voltage rating per plate is such that a single bridge consisting of 4 plates can be supplied from a transformer of up to about 18 V R.M.S. output. For higher transformer voltages the number of plates should be increased *pro rata*. Plates are made in a range of diameters, suitable for rectifying currents between about $\frac{1}{2}$ and 8 A. The output of bridge-connected rectifiers therefore ranges from 1 to 16 A. The specimen in *Plate 23.3* is a bridge-connected unit rated at 1 A, 18 V R.M.S. input.

SIGNAL DIODES

Germanium types

In germanium *signal* diodes connection is made to a block of N type of germanium and to a metal cats-whisker which presses on it. During manufacture a carefully controlled burst of heavy current is passed through the device which is found to convert the germanium in the region of the cats-whisker from N to P, thus forming the requisite junction (*Figure 23.3*).

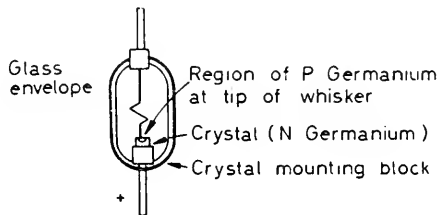


Figure 23.3

Diodes formed in this way will operate up to radio frequencies and have a P.I.V. ranging from 20 to 100 V. They can pass forward current up to about 50 mA. With one volt across them in the forward direction they mostly

NON-THERMIONIC DIODES

pass about 3 mA, though a few types pass 7 or 8. With 10 V across them in the backward direction the leakage current may be anything from a few microamperes to a milliamp. As a general rule, germanium diodes with a high P.I.V. have low reverse currents. There is also a slight tendency for diodes with very low forward resistances to have a poor P.I.V., but the reverse is not so; the STC 2X/106, for example, has a good P.I.V.—70 V—and passes a forward current of 7 mA at 1 V. On the whole the high P.I.V. diode is the better component.

Selenium type

These are among the smallest electronic components made, and are rather lower-power devices. The STC M1 and M3 diodes, for example, have a P.I.V. of 68 V but can pass mean forward currents of only 250 microamps and 1 milliamp respectively. The backward resistances are enormous—1,000 megohms for the M1 and 45 megohms for the M3—but the forward resistances are also rather high, 10 k and 1.2 k. M1 can be used at frequencies up to 5 M/c, but M3 only up to 100 kc/s.

Copper-oxide type

These have a P.I.V. of only 6 V per junction, and so have generally to be assembled into stacks. Copper-oxide signal diode stacks are made for a P.I.V. between 6 and 90 and for currents ranging from 100 μ A to 10 mA. They occupy a position intermediate between the germanium diode and tiny selenium M1 and M3 so far as forward current rating is concerned. The reverse current characteristic of copper-oxide rectifiers is rather worse than a good germanium diode, but the forward current reaches its rated maximum when the forward voltage is only 0.7. This, coupled with their stability of characteristics, makes them suitable for instrument (i.e. meter) applications.

Silicon junction type

These are not of the cats-whisker variety and are not suitable for use at frequencies above 100 kc/s or so. Their most important characteristic is a remarkable front-to-back ratio. Their forward resistance is only about 10 ohms, whilst the backward resistance is commonly hundreds of megohms. P.I.V.'s lie in the range 60–180 V and maximum mean forward currents are 100 mA. The excellent performance is combined with ability to operate—with reduced ratings—in ambient temperatures up to 150°C.

A group of comparable signal diodes is shown in *Plate 23.4*. No. 1 is germanium and whisker, P.I.V. 90, I_{\max} 50 mA. No. 2 is copper-copper-oxide, P.I.V. 72, I_{\max} 250 μ A. No. 3 is selenium iron, P.I.V. 68, I_{\max} 250 μ A. No. 4 is a silicon junction unit, P.I.V. 120 V, I_{\max} 100 mA.

VALVES

The newcomer to the field of electronics and radio is liable to bewilderment at the enormous range of valves which seem to be at his disposal; about 2,000 of them, in fact. He might be forgiven for feeling that to choose a valve in a particular case design calculations of exceeding nicety must be necessary. The author proposes to indicate how to master this *embarras de richesse*.

The reasons for the profusion of types are as follows:

(1) *Plurality of valve makers*. There are in Great Britain at least six major manufacturers who make ranges of valves which are largely overlapping, but who all have their own systems of designation. Thus manufacturers A, B and C may each make a valve with identical characteristics, but each will allot it a different type number.

(2) *Types of base*. There are extant at the moment 16 common types of valve base* and numerous more obscure ones. Partly this is explained by the progressive miniaturization of valves over the years, partly by differing countries of origin. Nevertheless it does not seem that the diversity is altogether rationally explicable.

(3) *Number of heater ratings*. Valves can be had with heaters intended for parallel operation from constant voltage supplies (1.4 V d.c., 2 V d.c., 4 V a.c., 6.3 V a.c., 12.6 V a.c.) or series operation from constant current supplies (0.1, 0.2, 0.3 amp a.c. or d.c.). This variety is attributable to the demands of the radio and television, rather than the general electronic, industry.

(4) *Composite valves such as triple-diode-triodes and triode-pentodes*. These are produced in response to specific needs in the radio and television industries and are only economic because they serve a mass market. In general electronic work it is a moot point whether composite valves are very helpful (apart from double triodes). They complicate stock-keeping and militate against flexibility in design.

By rejecting composite valves, adhering to one heater rating (6.3 volt a.c.), restricting oneself as far as possible to one base (the miniature 'Noval' or 'B9A') and one manufacturer it is possible to produce a 'short list' of rather over a dozen types. Between them they will do almost anything one is likely to want, and in time one gets to know their characteristics more or less by heart. The author's suggested short-list is as follows. Three kinds of base are used, and three manufacturers: they are mostly Mullard valves, but only because they happen to be kept by his local stockist.

* i.e., number of pins, their sizes, separations and dispositions.

VALVES

Maker	Type	Base	Description	g_m (mA/V)	r_a (k Ω)	μ	Remarks
Mullard	EB91	B7G	Double diode	—	—	—	PIV 420. I_{max} { mean, 9 mA { peak 54 mA.
Mullard	ECC81	Noval	Double triode	6.4	10	66	Pulse circuits
Mullard	ECC82	Noval	Double triode	2.2	7.7	17	Small power triodes
Mullard	ECC83	Noval	Double triode	1.6	62	100	Voltage amplifier
Mullard	EF80	Noval	Pentode	7.4	400	3,000	Pulse circuits.
Mullard	EF85	Noval	Variable μ Pentode	Up to 6	—	—	Where gain has to be controlled by a voltage.
Mullard	EF86	Noval	Pentode	1.8	2,500	4,500	Low noise, low hum, type.
Mazda	6F33	B7G	Pentode	—	—	—	Good suppressor control, for transi- strons, etc.
Mullard	EL37	Octal	25 watt power pentode	11	13.5	150	} Wattage figure is maximum power which can be dissipated at anode.
Mullard	EL84	Noval	12 watt power pentode	11.3	38	430	
Mullard	EL91	B7G	4 watt power pentode	2.6	130	340	
Mullard	EZ90	B7G	Full-wave rec- tifier, up to 70 mA	—	—	—	Can be heated from earthy 6.3 V supply to other valves if HT ∇ 450 V.
Mullard	EZ81	Noval	Full-wave rec- tifier up to 150 mA	—	—	—	Can be heated from earthy 6.3 V supply to other valves if HT ∇ 500 V.
Mullard	GZ33	Octal	Full-wave rec- tifier, up to 250 mA	—	—	—	5 V heater to be fed from separate winding in mains transformer.

A list of this kind is, of course, subject to continual revision as new types are announced.

In addition a few special types are necessary. The author suggests:

Description	Maker	Type	Remarks
Voltage stabilizers	Brimar	VR 75/30 VR 105/30 VR 150/30	75 V stabilizer 105 V stabilizer 150 V stabilizer } Can pass up to 40 mA.
Voltage reference tube	Mullard	85A2	(Octal) 85 V reference source (B7G)
Electrometer pentode to follow micro-electrode	Mullard	ME1400	Low grid current, otherwise a relation of the EF86. (Octal)
Tetrode thyratron	Mullard	2D21	(B7G)

SUNDRIES

SWITCHES

Switches fall into two clearly defined types, those for controlling the flow of signal currents and those for controlling the transfer of power. In the former the emphasis is on effecting good contact, by using noble metals for the contact faces—platinum, silver, gold—and by the employment of ‘wipe’. A wiping action in switches is one in which the contact faces are rubbed together as the switch is operated; in this manner the contacts are in some measure self cleaning.

Power control switches employ a spring toggle action which ensures that when the switch is opened the contacts are rapidly separated; in this way any arc which may be struck is quickly extinguished.

It is a mistake to try to use power switches for signal control, and vice versa. The thermal-capacity of the contacts of signal switches is low, and if these are used to break, for example, the primary circuit of a mains transformer the contacts are rapidly destroyed, even if one remembers to operate the component smartly. Similarly the use of power type toggle switches for signals—at least at low level—is also frequently disappointing. The contacts are of base metal, e.g. brass, and there is often no wipe, and in consequence the contacts are covered by an oxide film. Under the conditions of high electric field which obtain when these switches are properly used this film breaks down and current can flow. The feeble voltages constituting a signal do not affect the necessary breakdown and the result is a high resistance and noisy contact.

Despite the precautions taken, the performance of signal switches can eventually become unsatisfactory due to contamination of the contacts by air-borne dirt. The elegant solution is to employ sealed switches, but these are not yet common. The popular open variety may be cleaned by brushing the contacts with a grease-dissolving fluid such as trichlorethylene (Scroggie (*Radio Laboratory Handbook*) is against the use of commercial carbon tetrachloride, he points out that it frequently contains acid and may cause corrosion): a little lubricant may then be applied. Some typical switches are illustrated in *Plate 25.1*. No. 1 is a power-control toggle switch; No. 2 is a popular type of rotary signal switch, whilst No. 3 is a miniaturized and sealed version; No. 4 is a ‘key switch’ of the signal genus but, like the power toggle switch, capable of being operated—under conditions of experimenters’ duress—by the foot.

VALVEHOLDERS

The purposes of a valveholder are, fairly obviously, to retain the valve firmly in place, to make good connection to the valve pins and yet preserve a high resistance between them: modern valveholders are very reliable in all three

SUNDRIES

respects. The use of silver plated contacts is now general and the main difference between types lies in the insulating material used. This may be fibre, bakelite, nylon or PTFE, of which the first is the cheapest and least durable and the last the best. Bakelite or nylon are perfectly satisfactory for the general run of electrobiological apparatus. *Plate 25.2* shows nylon-insulated valveholders for the octal, miniature noval, and miniature-heptal-based valves which constitute the suggested short-list in the previous chapter.

KNOBS

The design of knob used for a particular control is by no means a subject unworthy of thought; correct choice here can make all the difference to the 'feel' of an instrument. For example, knob No. 1 in *Plate 25.3* is clearly for use with a continuously variable control where fine movements are necessary. The diameter is such that thumb and all four fingers may be arranged comfortably round it, and accurate operation secured thereby. Not so knob No. 2; this is for use with a rotary switch, where accurate indexing is performed by the switch itself. It can therefore be quite small, being grasped between forefinger and thumb and 'flicked' by an action of the wrist. Knob No. 3 may be regarded as a compromise for persons who like to lay out the front of an instrument with all the knobs the same, irrespective of function.

TAGBOARDS AND TAGSTRIPS

It is clear from the photographs of the components we have dealt with that whilst the larger items are provided with some fixing device involving bolts or nuts, the smaller rely on their own connections for fixing. At one time it was usual to employ the 'direct' method of wiring, in which such components were suspended by their own connections, but the appearance of the underside of apparatus built in this manner leaves much to be desired. The best that can be said for the method is that it ought to minimize stray capacitances; the worst is better left unuttered. Direct wiring is necessary at very high radio frequencies but is usually not justified below a megacycle. For the bulk of electrobiological apparatus a much neater procedure is to arrange small components in a row on one or more tagboards (*Plate 25.4*). Alternatively they may be soldered between two parallel tagstrips. Tagstrip mounting is more satisfactory for components which reach high operating temperatures, such as vitreous resistors, as the ventilation is better. The use of tagboards and tagstrips in equipment makes for much easier construction and subsequent servicing.

CONNECTORS

Complete electronic systems frequently comprise a number of discrete units connected together by flexible cables. Connectors are devices which enable these cables to be quickly and conveniently detached from the units in order that they may be separated where necessary. Although this may seem very obvious, it is easy to recall pieces of apparatus which cannot be moved either without much preliminary coiling, or else attendant trains of dragging cable.

CONNECTORS

As with knobs, choice of appropriate connectors merits a certain amount of thought. A selection of the important types is shown in *Plate 25.5*.

No. 1 is the least readily dismantled, requiring as it does the use of a screw-driver. It comprises a row of brass inserts, provided with grub screws, supported in an insulating plastic moulding. The required number of ways, up to a maximum of 12, can be cut off from stock with a pair of scissors. No. 2 is the familiar wander-plug and socket. Connection can be made extremely rapidly but an inadvertent tug on the wire can easily pull out the plug. A robust device is the screw terminal and spade; the scheme is suitable for high currents, of the order of amps, as the contact surface is large. No. 4 is a spring terminal; this has the firm grip of the screw terminal, but is quicker to use. No. 5 is a connector for a screened cable carrying signals; this particular one is for single core cables, but neat fittings may be had for twin- or 3-core cables. Nos. 6 and 7 are multi-way connectors for power supplies. In considering plugs and sockets for handling supplies where there are dangerous voltages, there is an important safety rule which determines whether the pins should be on the end of the cable and the sockets on the piece of apparatus, or vice versa. This is merely that 'power must always come out of the socket', not out of the plug; otherwise there is a risk of shock. Thus, if a power unit feeds an amplifier via a multicore cable, the power unit will have a fitting possessing sockets, whereas the amplifier fitting will be provided with pins. The cable will of course have pins in the power-pack end and sockets on the amplifier end.

No. 8 is a mains lead connector. 8A goes on the end of the lead, and 8B on the chassis of the apparatus. Notice that the safety rule is obeyed, but that the pins are recessed for neatness.

No. 9 is a 'jack' plug and socket. These may be used for similar purposes to number 5, but a more important use of jacks is their 'break-in' facility. Auxiliary contacts on the socket are operated by the action of inserting the plug. These may be wired up to achieve various things, of which perhaps the most important is that, as the plug is pushed home, a hitherto intact circuit is broken and the plug circuit is connected in, in series (*Figure 25.1*).

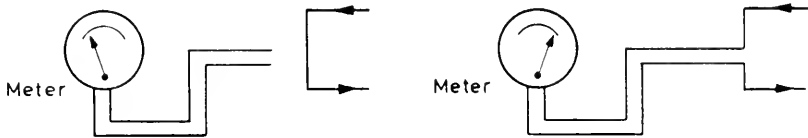


Figure 25.1

This property of jacks is extremely valuable in the metering of currents. To check a number of currents by having a separate meter for each is expensive and usually unnecessary. A better plan is to have one meter connected to a jack plug and break-in sockets in each of the circuits requiring measurement. One meter then serves for many. The same effect can, of course, be had by switching, but if the circuits in question are widely dispersed about the laboratory the extra wiring required is considerable; the jack system is usually preferable.

No. 10 is the familiar crocodile clip, in the standard and the miniature

SUNDRIES

sizes. A dozen or so lengths of insulated and flexible wire, supplied with a crocodile clip at each end, are extremely useful in electrophysiological work for effecting rapid *ad hoc* connections.

FUSES

It is good practice to provide the power packs of electronic apparatus with at least two fuses:

(1) A fuse in the mains transformer primary circuit. If this is not included and due to some fault condition the primary current becomes excessive, either the local house fuse blows—which is liable to annoy other people—or the local house fuse does not blow, in which case there is likelihood of fire. The coming of the fused mains plug removes the necessity for providing primary fuses on the apparatus itself.

(2) A fuse in the HT circuit, preferably in the connection between the transformer HT secondary winding and earth. This protects the transformer, rectifiers and smoothing choke against short circuits across the HT supply, particularly from faulty electrolytic capacitors.

Fuses are rated at 60 mA, 150 mA, 250 mA, 500 mA, 1 amp and 3 amp. In choosing an appropriate fuse rating for a particular application the steady current which normally flows may not be of much help because the fuse has to be able to withstand the 'switching-on surge'. Thus in the case of the HT fuse there is a surge to charge the smoothing capacitors which is greatest with non-thermionic rectification, intermediate with directly heated rectifier valves and small with indirectly heated types (that is, where the rectifier comes into action gradually). There is a surge through the primary fuse partly as reflected HT secondary surge current and partly as reflected LT secondary surge current. The latter occurs because, before the valve heaters have warmed up, their resistance is much lower than normally. In view of this, probably the best thing to do is to use the lowest fuse rating which does not persistently blow as soon as the equipment is switched on. Occasional fuse failures are then to be expected, and no fault need be suspected unless the replacement fails instantly also.

Plate 25.6 shows a neat fuse fitting which fits flush with the front panel of the apparatus.

INDICATOR LAMPS

These are useful for informing or reminding the user that circuits are alive. Small filament lamps are obtainable for low-voltage circuits (*Plate 25.7*) and diminutive neon bulbs may be employed for high tension voltages upwards of 80 or so. These bulbs do not possess built-in stabilizing resistors, and a small carbon component of a few hundred kilohms must be wired in series with the holder; otherwise, disaster is immediate.

SLOW-MOTION DRIVES

Where fine control is required of a variable resistor or capacitor and the use of a large knob proves inadequate, a slow-motion drive is indicated. A popular variety—an epicyclic gear—is shown in *Plate 25.8*. The velocity ratio obtained is about 4 : 1. Where extreme precision is called for, slow

CABLES

motion drives with ratios of 100 : 1 can be obtained. In checking over a particular drive before use it is essential to ensure that it is free from backlash.

CONNECTING WIRE

By this is meant material used to wire-up the inside of apparatus. Tinned copper wire insulated with PVC seems to be universal for wiring up electronic equipment nowadays, except that polythene insulation is better for EHT connections. Two procedures are clearly possible: to buy insulated wire and remove the insulation as necessary, or to buy bare wire—say 22 gauge—and slip on PVC 'sleeving' as required. For general work the writer prefers the former method, but the latter is quicker when half a dozen or so tags, rather close together, have to be connected.

Insulated wire may comprise a single core conductor of about 20 gauge, or may be 'flexible', i.e., stranded. Single conductor wire will take sharp bends, and equipment connected up with it, when the work is performed by an expert, looks extremely smart. In the hands of a novice the result is less happy, as every accidental bend to which the material is subjected leaves an ineradicable kink. The writer greatly prefers working with flexible wire which, if it does not make smart bends, at least takes up a course of graceful curves automatically.

Flexible wires are described by two figures separated by an oblique stroke. The first gives the number of strands, the second the diameter of strand in inches. The common flexible wires are:

	<i>Nominal current rating, amps</i>	<i>Resistance, Milliohms/yd</i>
7/0-0076	—	84
14/0-0076	2	42
23/0-0076	5	26
40/0-0076	10	14
70/0-0076	15	8·4

Of these, the first two are much the most important. 7/0-0076 is 'Instrument flex', only about 1 mm in diameter with insulation and, in virtue of its pliability, excellent for making lightweight connections to electrodes on animal preparations. It may also be used for general wiring work, but the insulation melts rather readily during soldering. Much easier to use is 14/0-0076, which, with insulation, is about 2 mm in diameter, and the writer strongly recommends this for the bulk of electronic wiring; for valve heater leads it is advisable to use a heavier gauge.

CABLES

Cables consist of one or more insulated flexible wires, with or without 'screening', i.e., a braided metal sheath which is earthed. Cables commonly have between one and seven cores, though types having two dozen or more are obtainable. The cores themselves are commonly 14/0-0076. The insulation may be PVC or polythene. If the cable is screened the screening may be

'overall', i.e., a single sheath enclosing all the cores, or each core may have its own screen. Thus the cable bringing the mains into the equipment might be twin core, screened overall, the object being to contain the electric field due to the large alternating voltage present and minimize the interference that such a field would cause if it were allowed to act on the input circuits of amplifiers. On the other hand if the conductors in a multicore cable are all carrying signals, or pulse waveforms, between which 'cross-talk'—due to mutual capacitance—must be eliminated, then the cores must be individually screened.

Choice of insulation depends on function. PVC is perfectly satisfactory for cables distributing power supplies *except* EHT, for which polythene is preferable. For carrying signals screened cable is usual and the limiting factor is then the capacitance between the core and the screen. Consider a typical case of a pre-amplifier coupled to a main amplifier by screened cable, and for simplicity's sake assume the circuitry is single-sided, so that the cable is single-cored. Let an upper turn-over frequency of 10 kc/s be acceptable and let the cable be driven from the anode of a pentode valve having a 100 k Ω load. Then $\omega_{\max} \approx 6 \times 10^4$ and the maximum permissible cable capacitance (stray capacitances being neglected) is

$$C = \frac{1}{\omega_{\max} R} \approx 160 \text{ pF}$$

Screened PVC-insulated 14/0-0076 has a capacitance of about 85 pF/ft., so the cable must be less than 2 ft. long. A polythene insulated cable of comparable overall thickness has a capacitance of 16.5 pF/ft., allowing a 9½ foot cable—a great improvement. It is possible to get very special cables having a capacitance of only 4 pF/ft.*

Some cables are shown in *Plate 25.9*. No. 1 is seven cores of 14/0-0076, individually screened, whilst No. 2 is similar, but 4 cored. No. 3 is 6 cored, 14/0-0076, screened overall. Nos. 4 and 5 are both single-cored screened polythene cables. It is important not to fall into the trap of supposing that the thicker cable has a lower capacitance because the core and screening are further apart. The capacitance depends on the ratio of the diameters of the core and screen, and in fact these are both 16.5 pF/ft. cables. The *raison d'être* for the thicker—and more expensive—cable is quite irrelevant to most electrobiological work; it is that the dielectric loss is lower at high (i.e. tens of megacycles) radio frequencies. No. 6 is a double-screened cable, useful for the input cable to cathodally screened cathode followers; the capacitance is 22½ pF/ft.

* Transradio C44.

FORMS OF MECHANICAL CONSTRUCTION

The 'traditional' form of mechanical construction for electronic equipment is the panel and chassis arrangement of the type shown in *Plate 26.1*. The controls are carried on the front panel with anything else which has to be visible or accessible from the front, i.e. dial lights, fuses and certain connectors, while all other items are carried on the chassis, with large components above the 'deck' and small ones below it. If the assembly is to be used alone it may be enclosed in a cabinet; if with a number of others a simple dust cover suffices and the group are carried in a vertical stack on a 'rack' (*Plate 26.2*). In order for this to be possible it is necessary that the panel width be standard, and until recently 19 in. was widely accepted as the norm. The panel height is not, of course, similarly restricted.

In recent years there have been considerable departures from this design. The progressive reduction in the size of components has led many engineers to feel that 19 in. is now too large, and a few firms have established a narrower panel, only 14 in. wide (*Plate 26.3*). Furthermore, attention is being increasingly focused on ready accessibility of components for servicing. With traditional construction, to gain access to the interior of a unit one has usually to: (a) break all connectors; (b) undo 4 bolts and withdraw the unit from the rack; and (c) undo further screws to remove the dust cover. If it is known exactly what modification or repair has to be carried out, one is now able to do it. More often there is an unidentified fault which requires that power supplies, etc., be re-connected to the unit in order that voltmeter and waveform checks may be carried out. Usually the various flexible connections do not reach unless the unit is offered up towards its original position on the rack and supported on a chair or steps, usually augmented by piles of books. Much thought has been devoted to producing alternative mechanical designs which simplify servicing procedures. Readers who are interested are referred to an excellent review by R. H. Garner (*Mechanical Design for Electronic Engineers*. London; Newnes).

In apparatus for scientific work the need to be able to get at the interior quickly seems to be even more vital than with the ordinary run of commercial gear. It is not always possible before embarking on the design of a new piece of equipment to specify exactly what it will be called upon to do. The time will surely come when some new demand will be made upon it, often at a few minutes notice during the course of an experiment. The writer therefore ventures to describe his own scheme for mechanical layout.

The apparatus is composed of small units measuring about 6 in. cube (*Plate 26.4*) which are rack- or console-mounted (*Plate 26.5*). By removing two thumbscrews the front cover may be drawn off a unit, revealing all valvholder and potentiometer connections, and allowing waveform checks and voltage measurements to be made to diagnose a fault. In this condition

FORMS OF MECHANICAL CONSTRUCTION

controls of the unit may still be operated as usual, since knobs and dials remain in place when the cover is removed. By pulling out the connectors and undoing two further thumbscrews, a unit may be pulled clear. Because the construction is 'inside-out' with respect to a conventional chassis, a wide solid-angle of access is possible to the connections of small components, facilitating speedy alterations or additions: also, and for the same reason, the units are very easy to make. There are no horizontal surfaces, and so ventilation by convection is good.

The unit in *Plate 26.6* contains 11 valve-envelopes (4 Dekatrons, 4 double-triodes, 3 trigger-triodes) and their associated circuitry, yet no difficulty was experienced in laying out the components within the available space.

TOOLS AND WORKSHOP FACILITIES

One of the reasons why small boys can bewilder their elders by producing home-made wireless sets is that electronic gear can be made with extremely humble tools and workshop facilities. Thus, whilst the home manufacture of quite a simple piece of mechanical engineering—a model steam-engine, perhaps—demands a proper workshop equipped with expensive items such as lathe, gauges, micrometer, etc., electronic apparatus of a kind can be built up in old tins with no other tools than a screwdriver, pliers, soldering iron, hammer and nails (the hammer and nails are used to punch holes as required). The workshop can be—and often is—the proverbial corner of the kitchen table.

On a rather higher plane, the research worker intent on making his own apparatus can produce professional-looking work by either: (1) buying ready-made racks, cabinets, panels and chassis (e.g. *Plate 26.3*), and cutting holes in them as required; (2) buying raw materials—sheet steel and aluminium, aluminium angle, etc.—and doing the whole job himself.

Plan No. 1 is advised for the beginner. If this procedure is adopted, all that is needed is a stout bench equipped with a vice having at least 3 in. jaws, and a box of hand tools; it is strongly recommended that they be his own personal property.

The author has reviewed his own tools and has prepared the following list. The items on it constitute the minimum with which he feels it possible to undertake any kind of electronic assembly work: the tools marked with an asterisk represent a short list with which it is possible to carry out work of a kind, but a high standard of workmanship cannot be expected with so impoverished an equipment.

6 in. rule, scriber, spring centre punch, divider and 2 ft. square for marking out.

Screwdrivers with $\frac{1}{8}$ "*, $\frac{1}{4}$ "* and $\frac{3}{8}$ " in. blades.

Nest of B.A. spanners*.

Files—flat, with safe edge, round, and half-round.

Bicycle spanner*, for potentiometer, rotary switch and toggle switch nuts.

Side cutters*.

'Q max' hole cutters: $\frac{5}{8}$ " in. for B7G valveholders; $\frac{3}{4}$ " in. for B9A valveholders; and $1\frac{1}{2}$ " in. for Octal valveholders.

Wheelbrace and drills: $\frac{3}{32}$ " in. (8 B.A. clearance); $\frac{1}{8}$ " in. (6 B.A. clearance)*; $\frac{5}{32}$ " in. (4 B.A. clearance)*; $\frac{3}{16}$ " in. (2 B.A. clearance); and $\frac{1}{4}$ " in. (0 B.A. clearance).

Handbrace and drills: $\frac{3}{8}$ " in. for making holes for potentiometers, rotary switches*; and $1\frac{1}{8}$ " in. 'Q max' cutter; $\frac{5}{16}$ " in. for making holes for $\frac{5}{8}$ " in. and $\frac{3}{4}$ " in. 'Q max' cutters; and $\frac{1}{2}$ " in. for toggle switches*.

Odd blocks of wood for supporting work when drilling.

TOOLS AND WORKSHOP FACILITIES

Snips, scissors, forceps, long-nose pliers—the last two are useful when wiring.

Soldering irons: 150 watt iron for tinplate work; 25 watt iron* for general wiring work. Miniature hacksaw.

Abrasive*—a hacksaw-like tool possessing a blade which cuts in any direction. Useful for odd-sized and odd-shaped holes.

In addition, one needs stocks of solder (flux-cored for preference), nuts and bolts in B.A. sizes—of which 4B.A. and 6B.A. are the most important—in $\frac{1}{4}$ and $\frac{1}{2}$ in. lengths, insulating tape, etc.

The difficulty about buying ready-made metal work is that it is expensive, that one has to use what one can get, and that the standard of workmanship and finish may be unnecessarily good where the construction of temporary or experimental apparatus is envisaged. Where a considerable amount of electronic work is foreseeable there is no question that plan No. 2 is more satisfactory; given the proper machines it is possible to turn out professional-looking sheet metalwork with very little practice. These machines are unfortunately rather large and far from portable, and it becomes necessary—unless one has access to someone else's—to have a definite workshop for them. The two most important are:

(1) *A 3 ft. guillotine (Plate 27.1)*. Sheet metal is supplied in pieces 8×4 ft. and 6×3 ft. It is important to remember to order the smaller size; nothing is more exasperating than, perhaps after considerable delay, to be supplied with material one cannot get into the machine.

(2) *A swing-beam folding machine*, for producing sharp, straight bends (Plate 27.2). The 'ordinary' type of folding machine is better than nothing, but is severely limited in the range of work it can perform.

If funds and space permit it is also worth considering:

A bench grinder so that worn or broken drills and screwdrivers may be repaired.

A bench, or pillar, drilling machine—these are really something of a luxury, but their use is unquestionably conducive to speed and accuracy.

A fly-press is also perhaps a luxury, but would show a great saving of time where large numbers of holes have to be punched. In addition, a fly-press can perform work not otherwise easily carried out, such as the forming of louvres.

Paint sprayer or crackling oven; the choice here depends on whether the smooth cellulose or the 'crackle' type of finish is preferred. Crackle finish is an excellent artifice for concealing the effects of indifferent workmanship below, such as surface scratches and improperly drilled holes. Cellulose is extremely difficult to brush on satisfactorily, but rather easy to spray. Crackle paint is normally brushed on, then the work is put into an oven for the crackling process to take place. Neither a paint sprayer nor oven are actually necessary, however. There is a proprietary type of crackle paint, Panl, which crackles up without oven treatment. As for cellulosing, the author must confess to an affection for the humble Flit gun, though it must be admitted that progress is rather slow.

PART III
TRANSDUCERS, ELECTRODES
AND INDICATORS

LIGHT SOURCES AND DETECTORS

F. W. CAMPBELL

The aim of this chapter is to give a brief outline of the main types of light source and photodetector which are readily available today. Recent improvements in manufacturing technique have placed a wide range of versatile light sources and detectors at the disposal of the worker who requires to utilize radiant energy in the laboratory. In some experiments, almost any light source combined with a simple detector will be sufficient but many techniques are only practicable if a careful choice of source and detector is made. In this survey emphasis will be placed upon the comparison and contrast of commercially available equipment by bringing together information from diverse sources. The theory of photoelectric phenomena is already well covered in the literature and will only be referred to briefly.

Photometric units

The reader who is unfamiliar with photometric units will find that reference to *Figure 28.1* will help to correlate his visual experience with the physical range of luminance commonly met. The dark-adapted human eye is a remarkably sensitive detector of visible light energy, only rarely bettered by physical detectors. Under certain conditions the eye can detect light entering at the rate of 500 quanta per second: at the other extreme, the brightest source normally encountered is the midday sun. Most practicable light sources are considerably less bright than the sun.

The eye is not equally sensitive to all wavelengths. If a monochromatic yellow and blue light of equal radiant energy were compared visually by a normal observer, the yellow light would appear much brighter. In *Figure 28.2* is plotted the variation of the reciprocal of the radiant energy required to create a fixed visual impression of brightness with the wavelength of the light observed. The ●—●—● curve is the average of many measurements made with light-adapted eyes (cone vision); the ○—○—○ curve is the average for dark-adapted eyes (rod vision). No photodetector has a spectral sensitivity curve like the human eye so that it is important to distinguish the sensitivity of the eye to a given light source from that of the particular photodetector. The photometric units in use are based on measurements made by the light-adapted human eye with normal colour vision.

The basic unit of luminous intensity is the candela (cd). It is of such a magnitude that the luminance of a full radiator (perfect black body) at the temperature at which platinum solidifies is 60 cd per cm². The unit of luminous flux is the lumen (lm). It is the flux emitted in a solid angle of one steradian by a uniform point source of 1 cd (*Figure 28.3*). Such a source will therefore emit a total of 4π lm, assuming it radiates equally in all directions.

LIGHT SOURCES AND DETECTORS

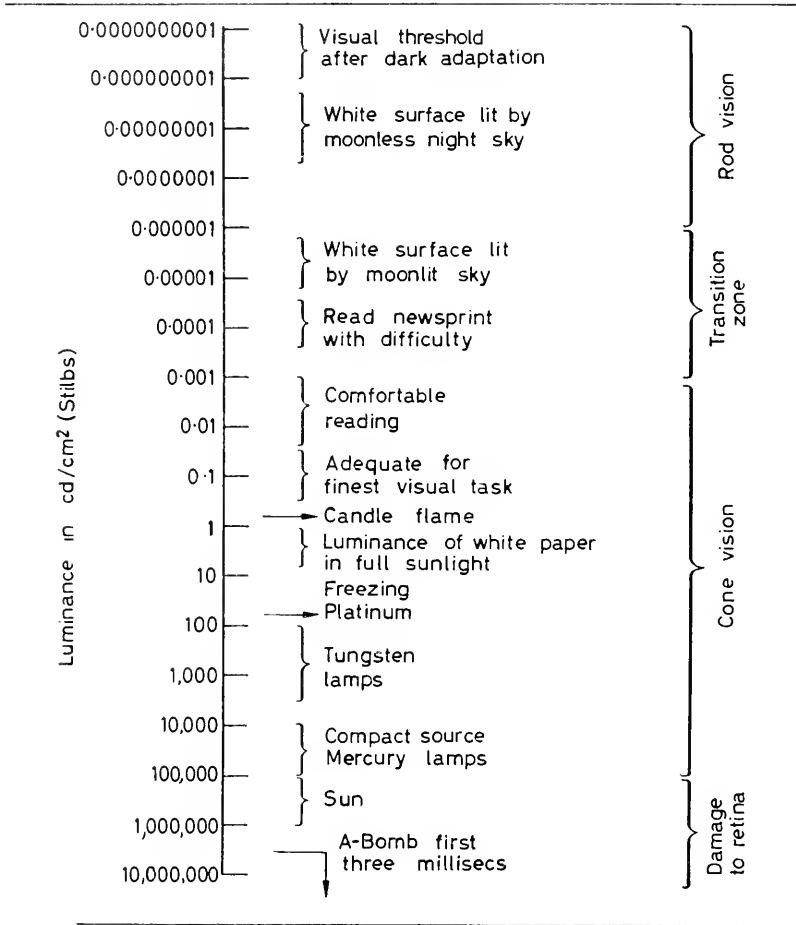


Figure 28.1

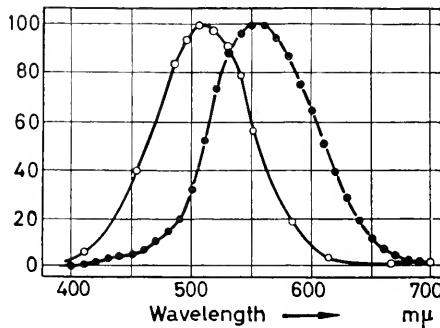


Figure 28.2 Luminosity curves for scotopic (rod) vision $\circ-\circ$ and photopic (cone) vision $\bullet-\bullet$.

LIGHT SOURCES AND DETECTORS

If a screen or photodetector is placed at a distance of 1 m from a source of luminous intensity 1 cd, the intensity of illumination falling on it will be one metre-candle or 1 lux. As the intensity of illumination varies inversely with the square of the distance from a point source, the illumination at any distance may readily be calculated (*Figure 28.3*). The foot-candle is sometimes used in Great Britain. One foot-candle (ft.-c) = 10.764 lux (lx) or metre-candles.

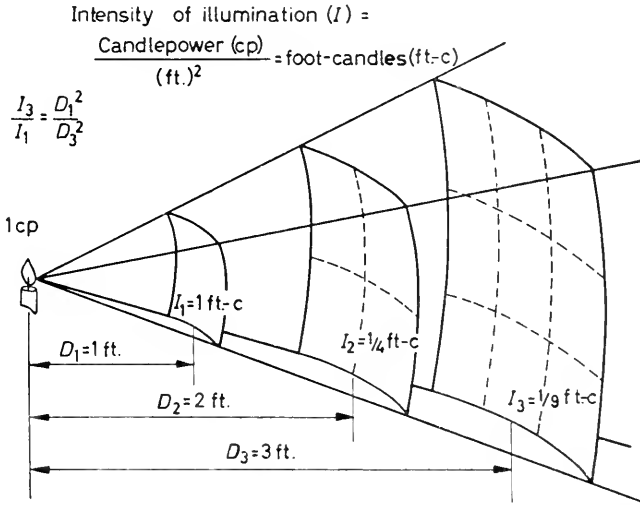


Figure 28.3 Diagram illustrating the inverse square law of illumination. The lumen is the luminous flux radiated by 1 candela into 1 unit solid angle

If a screen of area 1 m² is placed one m from a point source of 1 cd, it will intercept a quantity of light or luminous flux equal to 1 lm. If the screen has an area of 1 cm² it will intercept 1/10,000 of a lm, =10 millilumens (mlm). The brightness or luminance of a source of given luminous intensity will depend upon its area. One cd emitted from a source of area one cm² will have a luminance of one cd per cm² or one stilb. Thus the luminance of one cm² of solidifying platinum is 60 stilbs. The approximate luminances of other sources is shown in *Figure 28.1*. Unfortunately, there are a number of different units of luminance which tend to be used indiscriminately in the literature. A conversion table is given in *Table 1*. Stilbs (cd/cm²) should be used where possible. Consult Walsh¹ for full definitions of these units.

It is important not to confuse the concept of intensity of illumination with that of luminance. The illumination due to a source will decrease with the reciprocal of the square of the distance from it (assuming the source is small compared with the distance) but the luminance (or brightness) of the source is independent of distance. For example, the illumination produced at a point some distance from a 100 W pearl lamp will vary with the distance, but assuming a clear atmosphere the lamp will appear to have the same luminance to an observer at any distance, so long as the lamp subtends more than a few minutes of arc.

It is sometimes convenient to measure the intensity of illumination by

TABLE 1
Illumination Units

	lux (lx)	phots (ph.)	foot-candles (ft.-c)
Lux (metre-candles) (lm/m ²)	1	0.0001	0.0929
Phots (lm/cm ²)	10,000	1	929
Foot-candles (lm/ft ²)	10.764	0.001076	1

Luminance Units

	nits (nt.)	stilbs (sb.)	cd/in. ²	cd/ft. ²	millilamberts (mL)	foot-lamberts (ft.-L)	apostilbs (asb.)
Nits cd/m ²)	1	0.0001	0.000645	0.0929	0.3142	0.2919	3.1416
Stilbs (cd/cm ²)	10,000	1	6.452	929	3141.6	2919	31,416
Candelas per in ²	1550	0.155	1	144	486.9	452.4	4869
Candelas per ft ²	10.764	0.001076	0.00694	1	3.382	3.1416	33.82
Millilamberts	3.183	0.0003183	0.002054	0.2957	1	0.929	10
Foot-lamberts	3.426	0.0003426	0.002211	0.3183	1.0764	1	10.764
Apostilbs	0.3183	0.00003183	0.0002054	0.02957	0.1	0.0929	1

(Value in unit in left-hand column) × (conversion factor) = (value in unit shown at top of column)

LIGHT SOURCES AND DETECTORS

allowing a white surface to be lit by it, in turn measuring the luminance of the white surface. If a perfectly white screen which reflects 100 per cent of the light falling on it evenly in all directions is placed 1 m from a source of 1 cd it will have a luminance of 1 millilambert (mL) or 0.0003183 stilbs or 0.929 foot-lamberts (ft.-L) (see *Table 1* for other units). Freshly deposited magnesium oxide has a reflecting factor of about 98 per cent.

Radiation from incandescent solids

A black body or full radiator might be defined as a body which absorbs light of all frequencies and reflects none of the radiation which falls upon it. The amount of energy radiated by a perfect full radiator varies as the fourth power of its absolute temperature. If the temperature of the full radiator is known, the spectral distribution of its radiant energy can be calculated by Planck's equation. The spectral distribution of energy at various temperatures is shown in *Figure 28.4*. The radiation from a hot body becomes

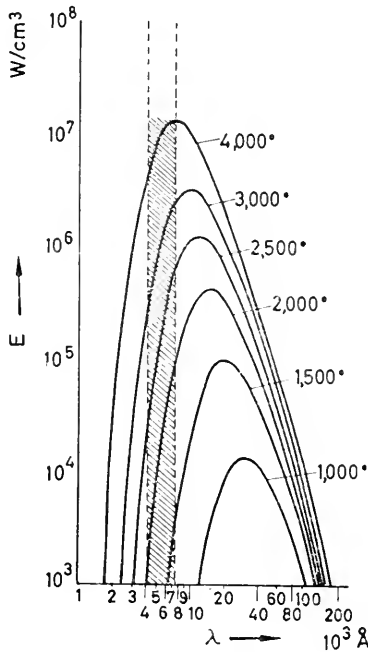


Figure 28.4 The spectral distribution of radiant energy emitted by an ideal black body at various absolute temperatures (after Sommer³). The shaded area represents the spectral limits of detection of the eye

visible to the human eye at about 1,000°K, where 1 part in 10^5 of the radiated energy is emitted in the visual range.

A perfect full radiator does not exist in nature, although in some instances the energy distribution from a source approximates closely to a full radiator in the visible region. For example, the energy distribution of tungsten in the visible region approximates closely to that of an ideal full radiator at a

LIGHT SOURCES AND DETECTORS

slightly higher temperature (*Figures 28.5 and 28.6*). The sun approximates to a full radiator at a temperature of 5,400°K, where about 40 per cent of the radiant energy is detected by the eye. In the case of tungsten filament

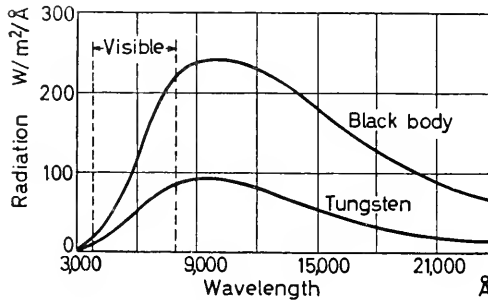


Figure 28.5 Comparison between the spectral distribution of the emission from an ideal black body and a tungsten surface, both at 2,870°K

TABLE 2

Colour Temperatures of Common Sources—from Spencer²

Light source	Colour temperature (°K)
Candle	1,900
Carbon filament lamp	2,080
Tungsten filament lamps 20 W	2,600
40	2,755
100	2,835
200	2,880
500	2,935
750	2,955
1,000	3,000
1,500	3,020
Photoparl lamp	3,100
Projection lamp, 500 W	3,190
G.E. (U.S.A.) 3,200°K lamps PS. 25	3,200
G.E. (U.S.A.) Photoflash S.M.	3,300
G.E.C. or B.T.H. Photoflood lamps Nos. 1 and 2	3,325
G.E. (U.S.A.) C.P. lamps	3,350
Mazda photoflood lamps Nos. 1, 2, and 4	3,425
G.E. (U.S.A.) Clear Photoflash	3,800
Sashalite	3,500–4,000
G.E. Daylight (blue) Photofloods (U.S.A.)	4,800
White flame carbon arc	5,000
Direct sunlight—early or late in day or winter	5,000
Sun's rays at sunset	2,000–4,000
Average noon sunlight	5,400
High intensity sun arc	5,500
Hazy sunlight, slightly overcast sky	5,800
G.E. (U.S.A.) Blue Photoflash	6,000
Sunlight plus light from clear sky at noon	6,500
Totally overcast sky	6,800
Light from a very hazy or smoky sky	7,500–8,400
Light from a clear blue sky	12,000–27,000

LIGHT SOURCES

lamps, the visual efficiency varies from 10 to 20 per cent, depending upon the filament temperature.

The actual temperature of a radiator is difficult to determine and it is more usual to measure its colour temperature (C.T.). This could be defined

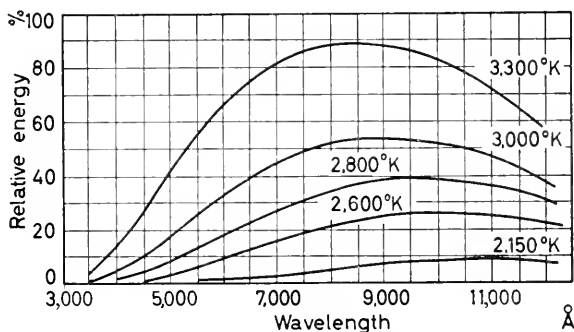


Figure 28.6 Spectral radiation of tungsten filament at various temperatures

as the temperature of a full radiator which would emit radiation of substantially the same spectral distribution in the visible region as the radiation from the light source, and which would have the same colour. Table 2 gives a list of colour temperatures of common sources.

LIGHT SOURCES

Characteristics of tungsten lamps

The diversity of form and size of the modern tungsten lamp makes it an indispensable light source in many experimental techniques. It is often possible to use a given lamp under conditions other than that intended by the designer, provided the basic characteristics of tungsten lamps are appreciated.

A modern tungsten lamp running at the high C.T. of about 3,400°K has a luminous efficiency of about 30 lm/W with a filament luminance of about 5,000 cds/cm². Its useful life under these conditions, however, would be short as it has to operate near the melting point of tungsten (3,650°K) where rapid evaporation of the tungsten occurs. General service lamps operate at lower colour temperatures—between 2,500 and 3,000°K—with a corresponding lower luminous efficiency, luminance (1,000 to 2,000 cd/cm²), and longer life (usually an average of 1,000 hours). The working characteristics of the various ranges of lamps are shown in Figure 28.7.

The useful life of a lamp is usually determined by the blackening of the bulb due to deposits of evaporated tungsten which gradually reduce the light output. This applies especially when a concave mirror is used behind a lamp, as light from the reflector system has to pass through the glass bulb three times. The blackening also increases the absorption of light by the glass to such an extent that the concomitant temperature rise leads to softening, and a 'blow out'.

For a given operating temperature, the greater the diameter of the filament the longer the life of the lamp. The amount of vaporization is proportional

LIGHT SOURCES AND DETECTORS

to the surface area of the filament and this increases linearly with filament diameter; but the cross-sectional area increases as the square of the diameter and thus increases more rapidly than the rate of evaporation. Thus low voltage lamps are more efficient than high voltage lamps of equal wattage, as they can operate at a higher temperature without increased evaporation.

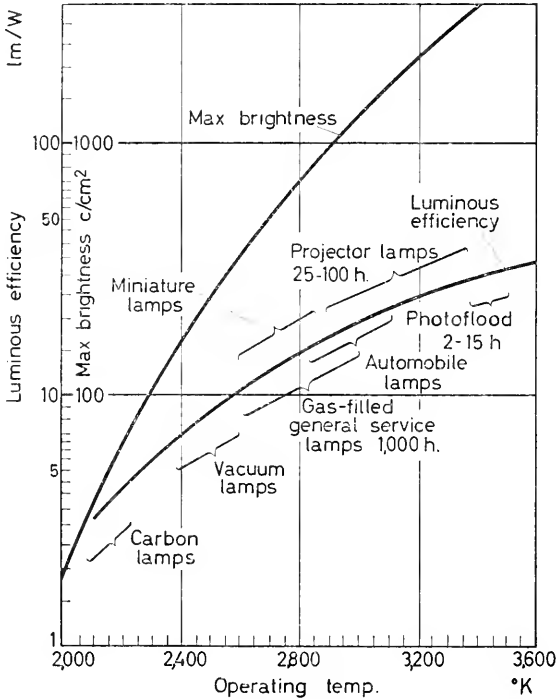


Figure 28.7 Working characteristics of various types of tungsten lamps: effect of filament temperature on efficiency and brightness (after Bourne³)

The life of a lamp varies approximately as the reciprocal of the fourteenth power of the voltage. Thus the life of a general service lamp would fall by some 13 per cent if its supply voltage were raised by 1 per cent. The effect of altering the supply voltage to such a lamp is shown in Figure 28.8. Approximate figures might be that for an increase of 1 per cent in voltage there would occur an increase of 0.5 per cent in current, 1.5 per cent in wattage, 3.6 per cent in light output and 2 per cent in light output per W.

Several useful points arise from these figures. First, if a bright compact light source is required it might be convenient to use a relatively cheap lamp such as a 6V automobile headlamp and over-run it by 10 to 20 per cent of its rated voltage. This would give the brightest source available from tungsten (5,000 cd/cm²) but with the disadvantage of frequent lamp changes. Lamp life could be slightly increased by bringing the filament slowly to its maximum temperature, perhaps by placing a suitable 'Brimistor'* in series.

* A resistor possessing a large negative temperature coefficient. Its resistance is initially high, but falls rapidly as the Brimistor warms up. ED.

LIGHT SOURCES

Secondly, as light output varies approximately as $V^{3.5}$, a well-regulated constant voltage is required if constant light output is desired. For low-voltage lamps lead accumulators are a convenient constant voltage supply if used on their 'voltage plateau', or if they are charged at a rate slightly less

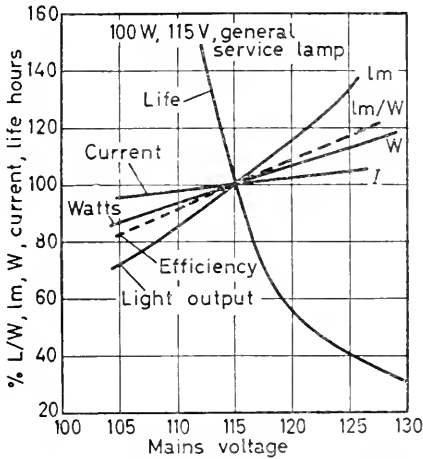


Figure 28.8 Variation of characteristics of tungsten filament lamp with mains voltage

than the discharge rate with the lamp running. Alternatively, the lamp may be supplied from the a.c. mains through a constant-voltage transformer of the saturated-core type, provided the 100 cycle ripple in the light output is not a disadvantage. The amount of ripple in the light output on a 50 cycle supply depends upon the thermal capacity of the filament. Approximate figures for a 200V, 100W service lamp would be 3 per cent, and for a 6V, 36W automobile headlamp, 1 per cent.

If a constant output is required from a lamp it is advisable to solder the supply leads directly on to the lamp base; there are few lamp-holders which can be relied upon invariably to produce a low resistance coupling. Leads to supply a meter to check the supply voltage should also be attached to the base; in this manner the meter reading is rendered independent of resistance changes in the supply leads. Failing this, it is better to control the lamp by placing an ammeter in series. However, even if the supply voltage to the lamp can be held constant a gradual decrease in light output will occur due to evaporation from the filament. This change is most rapid with a new lamp and some hours of aging should be given before setting in the equipment. The ideal way of ensuring a constant light output is to monitor it with a photoelectric cell and adjust the supply current either manually or automatically. Even with this technique, however, it will not be possible to hold the colour temperature in addition to the light output constant as the lamp ages.

Most modern tungsten lamps are filled with a mixture of 80 to 90 per cent argon and 10 to 20 per cent nitrogen to diminish the evaporation of tungsten and thus permit a higher operating temperature with a resulting improvement in light output per W. Argon is used on account of its low heat conductivity to minimize heat loss. The rarer and more expensive gas krypton is

sometimes used further to reduce heat loss and evaporation, with resulting gain in efficiency. Krypton is usually used in miners' lamps. These form useful light sources in portable equipment where heat and the weight of batteries may be limiting factors.

In a gas-filled lamp there is a layer of relatively stagnant gas around the filament. The thickness of this sheath is independent of the diameter of the filament. The effective cooling surface therefore decreases relatively as the diameter of the filament increases. The effective diameter of the filament may be increased by winding it into a coil and thus reduce heat loss. This process may be carried further with lamps having thin filaments, such as a 200 V, 100 W service lamp, by forming a coiled coil. Efficiency may be increased by 20 per cent with this technique.

Ultimately all tungsten lamps will fail due to fracture of their filament. Evaporation of the filament tends to be slightly uneven, so that some portion of the tungsten wire becomes thin, its resistance increases, and thus its temperature is raised locally to the melting point. Lamps running at a high colour temperature are particularly liable to fracture if subjected to vibration when alight. Automobile lamps and traction lamps are specially designed to withstand vibration by appropriate arrangement of the filament and its supports. Lamp failure due to filament fracture may be eliminated by running at a lower temperature and by switching on the power gradually. Higher voltage lamps with long filaments occasionally short-circuit on fracture due to the filament falling across one of the supports. To minimize the danger of a main fuse blowing under these circumstances, some manufacturers incorporate a local fuse in the base of the lamp.

Types of tungsten filament lamps

The range of lamps available today is very large indeed and it is usually necessary to consult the manufacturers' catalogues to ensure that the required type is in production. It will only be possible here to mention a few of the less familiar lamps which can be useful in the laboratory.

Class A1 projection lamps—This class of lamp uses a vertical grid filament to give a high light output in a direction normal to the plane of the filament. Projection lamps must be used vertically to prevent local overheating of the glass and 'blow out'. The C.T. is usually about 3,000°K and the life about 50 hours. They may be obtained in power ratings up to 1,000 W; the higher wattages require forced cooling.

These lamps are ideal for lantern and film projectors as an image of the filament grid may be focused in the plane of the slide or projection lens to fill it evenly with light (see Martin⁵, for optical details). The spaces separating the filaments in the grid can be brightened by forming an image of the filaments there by means of a concave mirror placed behind the lamp. This can almost double the light output in a forward direction. The heat output from these lamps can be very high and glass infra-red heat-filters may be required for some purposes. See Corning glass catalogue for details.

Class B1 floodlighting lamps—These lamps have a concentrated bunch filament with parallel coils arranged around the circumference of a cylinder and are for use with a parabolic reflector. They are mounted in round bulbs and may be used in any position except cap uppermost. They usually

have a colour temperature of about $2,800^{\circ}\text{K}$ and an average life of 800 hours. Their main use is for spotlighting.

Class E lamps—These are 500 W epidiastope lamps with a grid filament and a life of 100 hours.

Class F lamps—These are a most useful group of lamps for laboratory work. They are low voltage type for 6 or 12 V supplies with a heavy small filament, designed for microprojection work, etc., where a small uniform source is required. The solid ribbon filament type is particularly versatile (6 V, 18 A, 108 W, 50 hours life) and can be obtained with a vertical or horizontal filament. Because of the solid tungsten filament it is possible to avoid the uneven lighting which usually accompanies a coiled filament. The usual size of the filament is $20\ \mu$ thick, and about 2 mm wide \times 20 mm long. Philips Electrical Ltd. manufacture special versions of these lamps with optically ground glass or quartz front windows and reflection-free rear bulbs for precise photometric work.

A number of other small coiled-filament lamps down to 8 W are covered by this class. If over-run by 10 per cent of their rated voltage they form useful small bright sources.

Class G lamps—These are called Exciter lamps and are intended for sound film reproduction. They are manufactured to a high degree of physical accuracy to ensure interchangeability and to avoid the need for refocusing on refitting lamps. They are made in a number of low voltage ratings and wattages. The 10 V, 7.5 A, type are very convenient for over-running, as the evaporated tungsten deposits mainly at the top of the cylindrical bulb, leaving the filament area clear. Due to the heavy filament, the ripple or 'modulation' in the light output is small on a.c. supplies.

Automobile lamps—This type of lamp is too well known to require description. Automobile lamps can be obtained up to a rating of 48 W, with both transverse and axial filaments. The double filament types, each of 24 W, can be useful.

Unusual lamp types—It is not well known that small lamps of the flash lamp type can be obtained with a rating of only 6 V at either 40 or 60 mA (0.24 and 0.36 W). They are useful as indicators or fuses.

Thorn Electrical Industries make lamps with the very small dimensions of 6.3 mm bulb diameter and 14.6 mm overall length in the following ratings: 28 V, 0.04 A; 28 V, 0.08 A; 12 V, 0.1 A; 6 V, 0.1 A; 1.5 V, 0.75 A.

Miners' lamps, either argon or krypton filled, are available in a variety of sizes and ratings with both single and double filaments. These are valuable in portable equipment to save battery weight.

The largest low-voltage lamp available is an aircraft landing light at 26 V, 350 W.

For floodlighting a screen evenly to a high luminance reflector spotlights are useful. These lamps are parabolic in shape and have an internal coat of aluminium which forms an efficient internal reflector. Three types are available in the 150 W range with a polar distribution of output as shown in *Figure 28.9*. A 75 W version of the grid filament spotlight is also available with a peak candlepower at 0° of 1,250. A 275 W photoflood spotlight with internal reflector is made for photographic purposes running at a C.T. of $3,200^{\circ}\text{K}$ and having a life of a few hours.

LIGHT SOURCES AND DETECTORS

The 150 or 250 W infra-red spotlight with internal reflector designed for heating hatching chickens can often be usefully employed in the biological laboratory for maintaining the body temperature of anaesthetized animals.

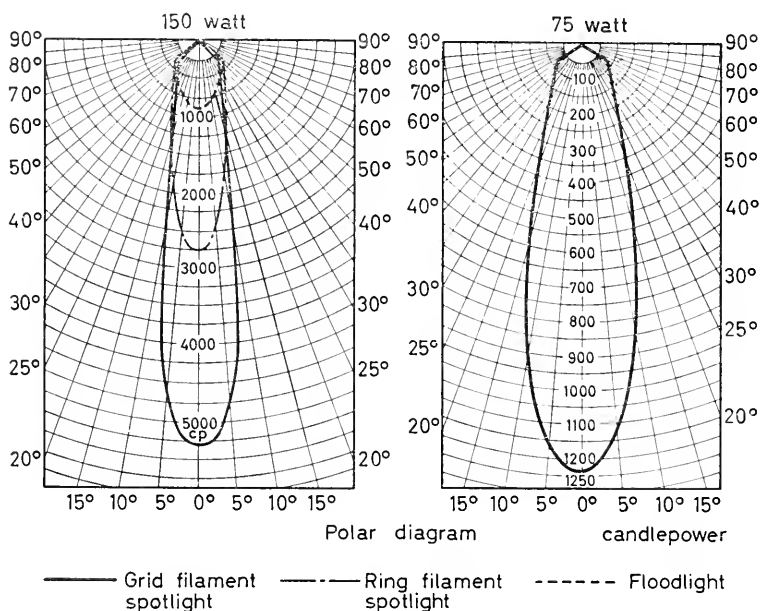


Figure 28.9 Polar diagrams of the candlepower distribution of internal reflector type lamps

Gas-discharge lamps

With a tungsten lamp running at a high temperature it is difficult to exceed a source luminance of 5,000 cd/cm². If a higher luminance is required it is necessary to use some form of arc discharge. The following sources are a few of the many available.

Type M.E. high-pressure mercury vapour lamp—This consists of a quartz bulb with two tungsten electrodes between which can be formed an arc in mercury vapour under high pressure (*Figure 28.10*). To protect the user against accidental rupture of the lamp it is enclosed either in a metal box with a tough glass window or in a glass envelope. It runs very hot but it should not be cooled by forced air as this reduces its efficiency. If mounted in a box with a concave reflecting mirror the effective luminance of the arc can be nearly doubled, for the arc is transparent to its own radiations and does not obscure the image formed by the mirror as is the case with a solid tungsten filament.

The arc, when it strikes, acts as a negative resistance so that appropriate supply circuits are required. It will operate on either an a.c. or d.c. supply of 200 V or more, using a choke and resistance to limit the supply current and to provide an inductive starting voltage surge. It takes about 10 minutes for the lamp to reach its maximum light output after starting: if it is then extinguished it will not re-strike for another 10 minutes until it has cooled.

LIGHT SOURCES

The life of the lamp varies inversely with the number of starts to which it is subjected, but with normal use a life of 500 hours can be expected, by which time the light output will have dropped to 80 per cent of its initial value.

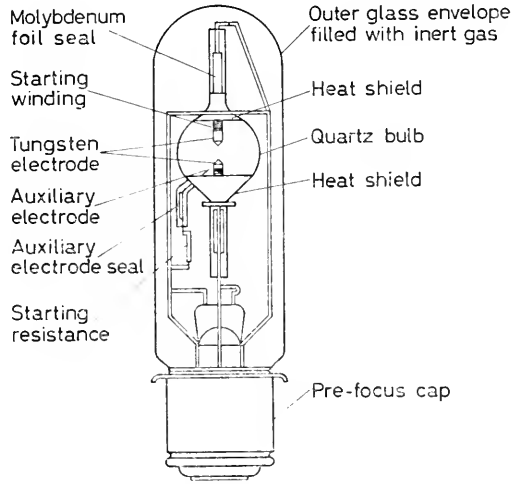


Figure 28.10 Construction of Mazda type ME compact source lamp

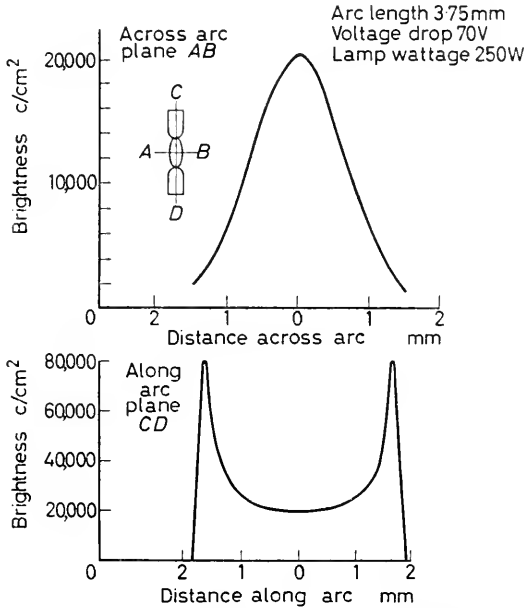


Figure 28.11 Brightness distribution of arc of 250W type ME lamp

The luminance varies from point to point on the arc (Figure 28.11). Most of the arc area has a luminance of 20,000 cd/cm^2 , but there are two small areas, near the electrodes, which reach a luminance of 80,000 cd/cm^2

LIGHT SOURCES AND DETECTORS

when operating on an a.c. supply. On a d.c. supply there is only one bright area—at the cathode.

Figure 28.12 shows how a change in power supply voltage affects the

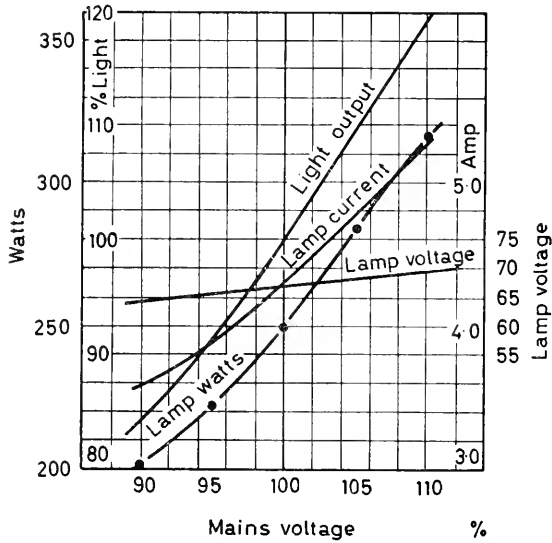


Figure 28.12 Effect of variation of mains voltage on characteristics of 250 W type ME lamp. Lamp operating from B.T.H. MRA 246 choke on a.c. supply at 50 c/s

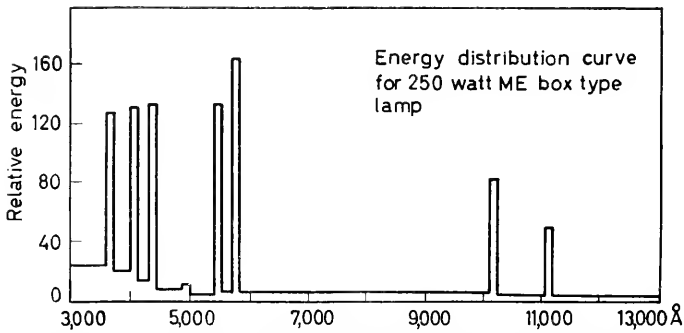


Figure 28.13 Energy distribution curve for 250 W type ME high-pressure mercury vapour lamp

lamp characteristics. A rise of 10 per cent in voltage increases the light output by 20 per cent, compared with 36 per cent for a 250 W tungsten projector lamp. The life of the compact source lamp is only slightly affected by supply voltage changes and there is also very little change in the spectral distribution of the light output with change of supply voltage.

The spectral energy distribution curve for a 250 W ME lamp is shown in Figure 28.13. At a supply frequency of 50 cycles the modulation of light output causes it to fluctuate about 60 per cent above and below the mean level. Lamps of this type may be obtained with a variety of ratings up to

LIGHT SOURCES

5 kW. The larger lamps do not have brighter arcs, but simply bigger arc areas.

The strong emission line at $365\text{ m}\mu$ in the ultraviolet with mercury discharge lamps is frequently used for inducing fluorescence in certain materials. By using a Wood's glass filter this radiation can be isolated from the visible spectrum, and also from the shorter wavelength emissions at 313 and $334\text{ m}\mu$, which can cause a dangerous and painful conjunctivitis of the eyes, even with short exposures. The 125 W Type MB w/v lamp is a mercury vapour discharge lamp with a Wood's glass bulb around it. It can be used close to the skin and eyes without danger. Philips Electrical Ltd. now manufacture a useful 40 W tubular mercury lamp with a special dark blue glass which is almost opaque to visible radiation (*Figure 28.14*)

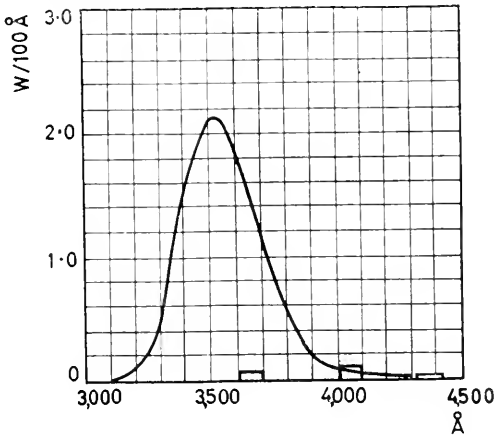


Figure 28.14 Energy distribution curve for a Philips' TL Blacklight Blue fluorescent lamp

for fluorescent studies. If all the visible and shorter ultraviolet radiations are required a useful small ultraviolet lamp is the Philips OZ 4 W which requires a 24 V, 0.35 A supply, the lamp itself dissipating about 4 W. It can be used for generating ozone, as the special glass bulb is transparent to the shorter ultraviolet radiations. In consequence the eyes must be screened.

Table 3 gives the main spectral lines in mercury vapour discharge lamps,

TABLE 3
The Main Spectra Emission Lines From a Mercury Vapour Discharge Lamp With Suitable Isolation Colour Filters

<i>Mercury vapour line in $\text{m}\mu$</i>	<i>Wratten filter No.</i>	<i>Ilford filter No.</i>
577, 570	22	808
546	77 or 77A	807
436	50	} 805 + 806 + 809
365	18A	

LIGHT SOURCES AND DETECTORS

with some filters which may be used for isolating these lines. The relative intensities of these emission lines depend upon the pressure of the mercury vapour, the design of the electrodes, and the nature of the quartz and glass around the arc; factors which naturally vary from one lamp to another. The higher pressure lamps also have quite a high-intensity continuous spectrum.

Ferranti Ltd. manufacture three-electrode gas-filled lamps filled with either helium, neon or mercury-argon, designated Types 511, 5521 and 5531 respectively, which provide a wide variety of spectral lines at moderate luminance. The light intensity is approximately proportional to the anode current.

Xenon arc lamps

The high-pressure compact-source xenon arc lamp is a powerful producer of radiation, extending from short ultraviolet at $200\text{ m}\mu$, through the visible spectrum to a peak at $900\text{ m}\mu$, in the near infra-red, and extending to about $2,000\text{ m}\mu$ in the infra-red spectrum (*Figure 28.15*). Unlike the mercury

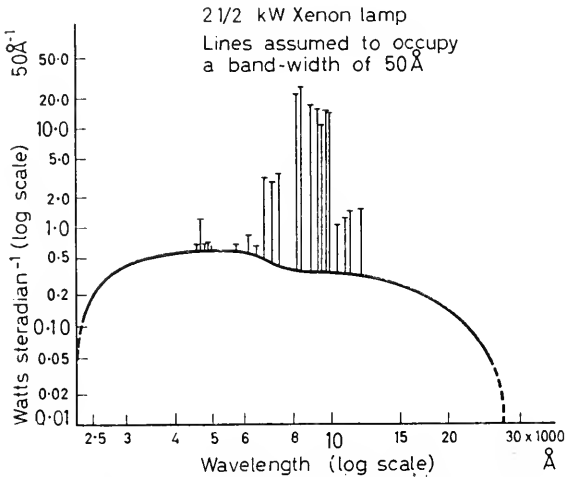


Figure 28.15 Energy distribution curve for a xenon high-pressure discharge lamp

high-pressure arc, it radiates richly at the red end of the spectrum. This good colour balance makes it particularly suitable for the viewing of coloured objects.

A wide variety of the xenon arc lamp is made. The usual construction is to seal two tungsten electrodes in a quartz mount and bulb, with a xenon gas pressure of about 2 atmospheres when cold. The pressure rises greatly when the lamp temperature rises, the strength of the quartz bulb being the limiting factor in the design. A typical small lamp is the British Thomson-Houston type E2 with an operating current of 25 A, arc voltage of 17, arc wattage 375, luminous efficiency 23 lm/W. The arc length is 6 mm with a centre brightness of $3,500\text{ cd/cm}^2$ and two bright spots near the electrodes of $10,000\text{ cd/cm}^2$ on a.c. operation. Lamps are also available

LIGHT SOURCES

TABLE 4
A List of Some Typical Xenon Electronic Flash Lamps at Present Available

No.	Dimensions mm		Max. ratings Watt/sec	Approx. minimum trigger voltage kV	Operating voltage		Rate of flash at maximum dissipation	Flash duration (micro-sec)		Capacitance at normal operating Volts/ μ F	Base
	Length excluding pins	Diam of glass			Max	Min		above $\frac{1}{3}$ peak	above $\frac{1}{2}$ peak		
FA1	150 \pm 7	64 \pm 4	1,000	8	2,700	2,000	1 in 10 sec	1,250	600	320	3 pin special
FA2	90 \pm 5	46 \pm 2	500	8	2,700	2,000	1 in 10 sec	870	300	160	3 pin 5 amp
FA2s	90 \pm 5	46 \pm 2	—	8	2,500	1,500	300 per sec	—	—	—	3 pin 5 amp
FA4	70 mm inside diam of circle		300	3	2,700	2,000	1 in 10 sec	440	180	96	Flexible leads
FA5	148 \pm 4 (overall)	32 max	150	10	2,000	1,000	1 in 10 sec	250	65	—	0-3954 Ferrules
FA6	70 \pm 2	31 \pm 2	100	3	1,100	500	1 in 10 sec	410	180	200	UX 4 pin
FA7	80 \pm 2	31 \pm 2	100-200	3	2,700	2,000	1 in 10 sec	800	290	64	UX 4 pin
FA7s	80 \pm 2	31 \pm 2	—	3	2,500	1,500	300 per sec	—	—	—	UX 4 pin
FA8	55 \pm 2	25 \pm 1	75	4-5	500	250	1 in 10 sec	800	270	600	Internat. octal
FA9	80 \pm 2	31 \pm 2	200	3-5	1,100	500	1 in 10 sec	2,000	770	400	UX 4 pin
FA10	50 \pm 2 (overall)	—	100	3-5	270	200	1 in 10 sec	2,450	930	2,750	2 pin 2 amp
FA12	50 \pm 2 (overall)	—	50	3-5	200	145	1 in 10 sec	3,200	1,200	3,100	2 pin 2 amp

with electrodes designed for d.c. operation. A high-voltage starting circuit is required, with suitable chokes and resistances for steady running. Average life is 500 h. The lamp starts and reaches full luminance immediately. It may also be restarted instantly when hot.

Another version of the xenon lamp is type FA5, which has an extra electrode for pulse starting. When running continuously it takes 10 A at 15 V. This 150 W lamp has a maximum luminance of 1,000 cd/cm² with an arc width of 2.5 mm. Whilst burning, a high current pulse may be applied of some 300 A when the arc becomes a sphere of 6 mm diameter.

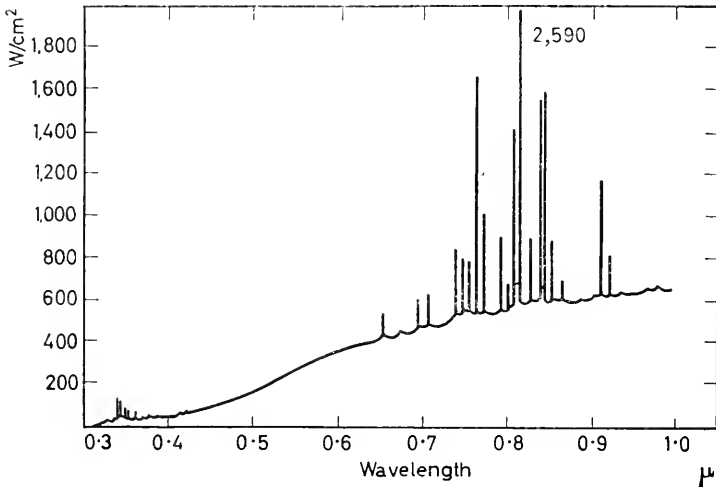


Figure 28.16 Spectral distribution of radiation from a 100 W zirconium concentrated arc lamp (Sylvania Electric Products Inc.)

A 200 joule capacitor-discharge may be used. The peak luminance of the arc reaches about 100,000 cd/cm². Refer to Beeson⁶ for circuit details. The lamp may also be used as a compact stroboscopic source with a mean rating of 125 W at a repetition frequency exceeding 160 flashes per second.

The requirement for intense short duration light sources in photography has led to the development of many different types of single-flash and multiple-flash stroboscopic, xenon lamps. These will be familiar to readers and are well described in the photographic literature. Table 4 summarizes the properties and power requirements of a few of these.

The zirconium arc lamp—Sylvania Electric Products Inc., New York, manufacture a variety of high-intensity zirconium arc lamps. Two electrodes are sealed into a glass bulb filled with argon. One acts as an anode, the other is a specially prepared refractory oxide cathode. When the arc strikes, the oxide surface is raised to its melting temperature and molten zirconium is liberated and vaporized. The vapour is drawn to the cathode, renewing the surface and ensuring a long life. The light is emitted from the molten zirconium, and its vapour in the immediate vicinity of the cathode.

The smallest lamp (2 W) has a source of about 0.08 mm diameter with an average luminance of 7,300 cd/cm² and the largest lamp (300 W) has

LIGHT SOURCES

a source of about 3 mm diameter with an average luminance of 3,700 cd/cm². As the light originates from a point on the cathode surface it is emitted in one hemisphere only and has a cosine type of distribution. The spectral distribution of the radiation from a 100 W lamp is shown in *Figure 28.16*.

About 1 to 2 kV are required for starting and, except for the 2 W version, at least 50 V d.c. for running. The 2 W version requires 200 V. Starting circuits are available from the makers.

General remarks on arc lamps—One of the main disadvantages of all the types of discharge lamp described above is instability in the position of the arc. While modern design of electrodes has done much to diminish the tendency of the arc to 'wander' in relation to the electrodes, all of them do so to some extent, especially as the lamp ages. There is usually an optimum current at which the arc has the greatest stability and the manufacturers' recommendations should be adopted. If absolute stability of position or luminance is required the tungsten lamp is at a definite advantage.

Linear light sources

It is sometimes necessary to control the intensity of a light source with great precision, both in luminance and in time, so that a light output of a particular waveform is achieved. A number of 'glow modulator tubes', such as are used for sound film recording, are available.

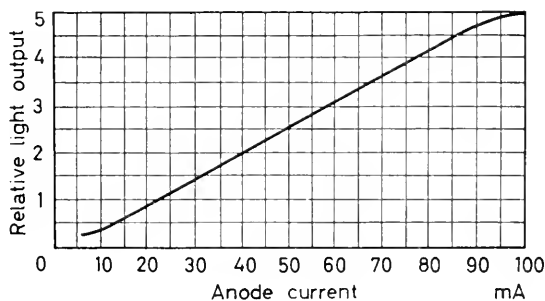


Figure 28.17 Relation of light output to anode current of a Ferranti type CL40 linear light source

Ferranti linear light sources CL40 and CL41 are good examples. These are valves with heater, cathode, anode and trigger and are filled with mercury-argon gas. The light output is proportional to the anode current over the range from 10 to 90 mA and it may be modulated at frequencies up to 12 kc/s (*Figure 28.17*). The CL40 has a circular light source of 3.2 mm diameter and CL41 a source 3×0.5 mm. A recommended circuit is shown in *Figure 28.18* where the modulator tube is connected in series with a hard valve which can pass 100 mA. The potentiometer *R4* is present to limit the current to 100 mA when *R5* is at minimum resistance. *R5* is then adjusted so that the tube is operating at the required mean current. Anode voltage should not be applied until the filaments have run for at least 5 minutes. The CL50 is filled with mercury/argon and the CL51 with helium. These tubes have cylindrical light sources measuring 5×32 mm,

LIGHT SOURCES AND DETECTORS

broadside view. A close approach to the light source is possible. Again the light output is proportional to the anode current.

For many purposes, such as time marking on slow moving photographic film, a simple neon lamp such as the Hivac CC8L will suffice, although the luminance of this source may not be high enough for fast time-marking or for vision experiments.

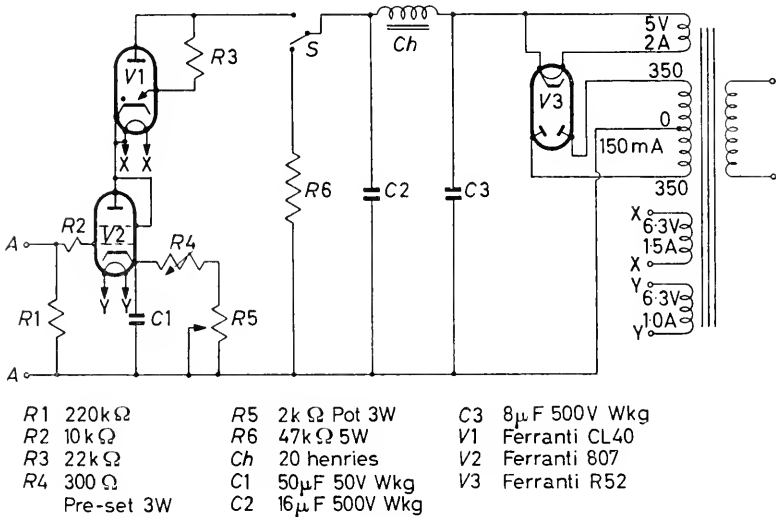


Figure 28.18 A recommended circuit for controlling a CL40 linear light source

Cathode-ray oscilloscope sources

The widespread use of the cathode ray oscilloscope in the modern laboratory often provides a convenient opportunity to use the C.R.O. as a modulated light source for experimental purposes. It offers the advantages of slow or fast adjustment of the position of the source, of adjustable source size and, of course, precise modulation of the intensity.

A number of phosphors are now in use for cathode ray tubes. The spectral energy distribution of the type B screen (Mullard Ltd.) is shown in Figure 28.19. It is the screen usually used when photographic records are required. Its spectral output matches very well the type S4 photocathode used in many photocells, e.g. the 931 A photomultiplier tube. It has a short duration phosphorescence, the light output dropping from 100 to 10 per cent in 3 msec and to 1 per cent in 9 msec. Type F and L screens are green in colour and are preferable for visual monitoring. They have a moderately long afterglow time, the light output falling from 100 to 10 per cent of luminance in about 20 seconds. Type W is used for television screens and is a mixture of several phosphors to give a near-white visual impression. Type R is a long-persistence screen developed for radar displays, but useful in the biological laboratory for studying slow non-repetitive events such as the electrocardiogram (Figure 28.20).

LIGHT SOURCES

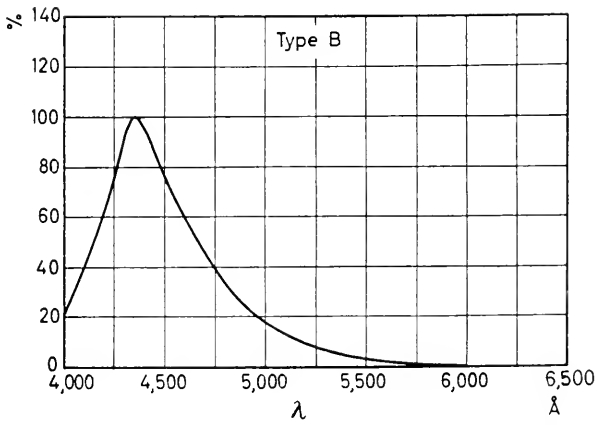


Figure 28.19 Relative spectral energy distribution curve for a Mullard type B luminescent screen

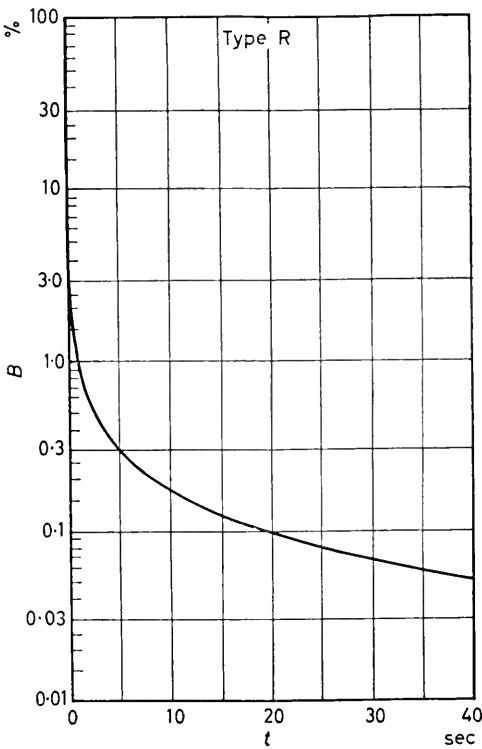


Figure 28.20 Persistence characteristic curve for Mullard type R luminescent screen

For flying-spot scanning techniques the MC 13-16 type of tube has been specially developed. This has a blue-violet screen well matched to type S4 photocathodes. The luminance is reduced to 36 per cent of the initial peak value $<0.1 \mu\text{sec}$ after the excitation is removed and this permits very high resolution and scanning frequencies. It has a useful screen diameter of 108 mm.

Ferranti Ltd. have developed a number of small grid-controlled triode tubes with a fluorescent screen designed to produce single or trains of light flashes of high luminous intensity. The four types differ only in the screen phosphor, as follows: CL60, green screen, decay time to 36 per cent less than $1 \mu\text{sec}$; CL61 blue screen, less than $3 \mu\text{sec}$; CL62, ultraviolet screen, $0.1 \mu\text{sec}$; CL63, yellow-green screen, $6 \mu\text{sec}$. The unfocused luminous area is 2 in. in diameter but it may be reduced to $\frac{1}{4}$ in. diameter by a simple magnetic focusing coil. The light output of types CL60 and CL61 is 3,000 cd approximately and of type CL63 7,000 cd. A current of $100 \mu\text{A}$ at an anode voltage of 20 kV is required. The anode may be pulsed up to 50 kV at 100 mA at infrequent intervals, or short duration trains of pulses may be given. A flash duration of less than $1 \mu\text{sec}$ can be obtained with the CL60 tube.

Electroluminescent sources

A phosphor is a substance which absorbs energy from an exciting source and converts part of it into light. The primary excitant might either be a photon, say of ultraviolet light as in a fluorescent mercury-vapour discharge tube, or a charged particle such as an electron in the case of a cathode ray tube, or an α -particle in the case of a scintillation counter. It is not well known, however, that a phosphor may also be excited by placing it in an alternating electric field. The phenomenon is known as electroluminescence. Early investigations in 1920 produced cells whose light output was rather feeble, but since 1950 increased research effort has resulted in the production of phosphors with a useful luminance.

The structure of an early electroluminescent 'panel' is shown in *Figure 28.21a*. A more practical example is shown diagrammatically in *Figure 28.21b*. Ballantyne⁷ and Bowtell⁸ should be referred to for details of construction.

The light output of these panels depends upon the applied voltage, as shown in *Figure 28.22*. The maximum voltage which may be used depends upon the strength of the dielectric layer and is usually about 1,500 in a well-made panel. Unfortunately, at the usual domestic supply voltage, 200 to 250 V, the luminance is not great and a step-up transformer is required. Bowtell⁸ describes a green, zinc sulphide panel which has a luminance of 2 ft.-L on a 50 c/s, 240 V, supply and a luminance of 50 ft.-L on a 500 c/s, 600 V, supply. White panels have a luminance of about half this value.

The applied voltage must be alternating, and frequency affects the performance of the panel. The luminance increases linearly with frequency, reaching a maximum at frequencies of the order of a few kc/s. With multiple phosphors the colour of the resulting light output may change with frequency as the emission of each phosphor in the mixture will alter differentially with change of frequency. Fortunately, the light output of these panels is satisfactory at 400 c/s, and this makes them particularly useful as signboards in aircraft where supplies at this frequency are generally available.

LIGHT SOURCES

Some attempts have been made to produce a light amplifier, using this phenomenon, by placing a light-sensitive photoconductive layer between the conducting glass and the phosphor dielectric layer, but the efficiency of such

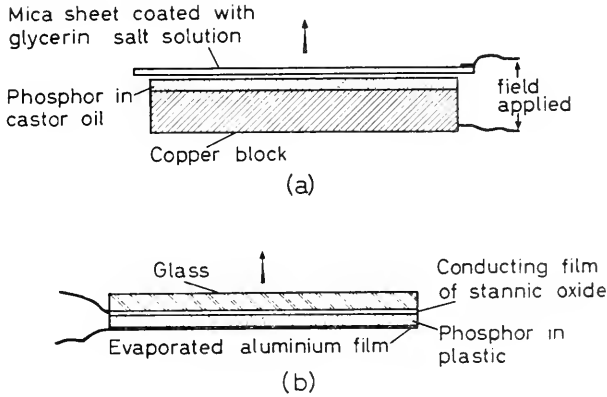


Figure 28.21(a) Destrian's original electroluminescent cell;
 (b) A conventional electroluminescent light source

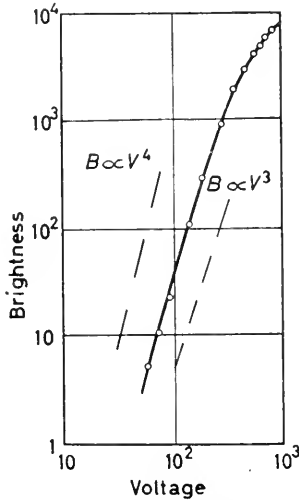


Figure 28.22 The variation of the mean brightness of an electroluminescent panel with voltage

amplifiers has so far been poor. Further development of suitable phosphors and rapidly acting photoconductors may improve their performance and also extend their use to infra-red and ultraviolet image-converters. If combined with photoconducting cadmium sulphide an electroluminescent X-ray intensifier may be practicable.

A bi-stable device using optical feedback has also been described, using a green-sensitive cadmium-sulphide photoconductive crystal and a green

LIGHT SOURCES AND DETECTORS

electroluminescent layer. Assemblies of such small elements could be envisaged for use in counter and computer circuits (Kazan and Nicoll⁹).

PHOTOELECTRIC DETECTORS

To measure the efficiency of a light detector it is necessary to place before it a light source of known intensity and measure the resulting increase in current flowing through the detector. The amount of radiant energy the detector will collect will of course depend not only upon the intensity of the light source but also upon the distance of the cell and the area of the light-sensitive element. The number of lumens (lm) incident upon the photo-sensitive element is given by the equation:

$$\text{lm} = \frac{C \times A}{d^2}$$

where C = intensity of the source in cd, A = area of photosensitive element in cm^2 , and d = distance of the source from the photocell in cm.

If a cell is so calibrated, and if it has a spectral response curve identical with that for the eye, the current output per lm will remain independent of the precise spectral distribution of the light source whether it be a full radiator at any colour temperature or a monochromatic coloured light: but no such photodetector exists. It is therefore necessary to standardize the spectral nature of the light source used for calibration in order that different photocells may be compared. In Great Britain a tungsten lamp at a colour temperature of $2,848^\circ\text{K}$ is usually used, while in America one at a C.T. of $2,870^\circ\text{K}$ is preferred. At the latter temperature the wavelength at which maximum power is radiated is almost exactly 1μ .

Photoemissive type of detector

The simple well known type of photoemissive cell can be obtained in diverse sizes and shapes, either vacuum or gas filled, and with three main types of photocathode. This kind of cell consists of an electrode on which is deposited a special layer of some metal such as caesium or antimony. When radiant energy of suitable wavelength strikes this layer electrons are liberated which can be attracted to, and collected by, an anode maintained at a positive potential. Any change in the current flowing through the cell can be detected by placing a galvanometer in series or by measuring the potential changes occurring across a load resistor.

The spectral sensitivity of the three main types of photocathode commercially available is shown in *Figure 28.23*.

Type A cathode—This is called type S4 in the U.S.A. It is an antimony-caesium cathode enclosed in a lime glass envelope. Its peak sensitivity is at the violet end of the spectrum, where the quantum efficiency of a good cell may reach 20 per cent. That is, on the average, 1 in 5 of the light quanta striking the photocathode are absorbed and influence the flow of electrons. A similar surface enclosed in ultraviolet transmitting glass is called type S5. It has a high sensitivity near the mercury resonance line at $253.7 \text{ m}\mu$.

Type B cathode—This is called type S8 in the U.S.A. It is a bismuth-oxygen-antimony-caesium surface. Its spectral sensitivity extends further

PHOTOELECTRIC DETECTORS

into the red than type A, which has little response beyond 630 $m\mu$. However, it is not so sensitive to the shorter wavelength.

Type S cathode—Type S1 in the U.S.A.: the overall sensitivity is much less than types A or B but it has the great advantage of being usable out to

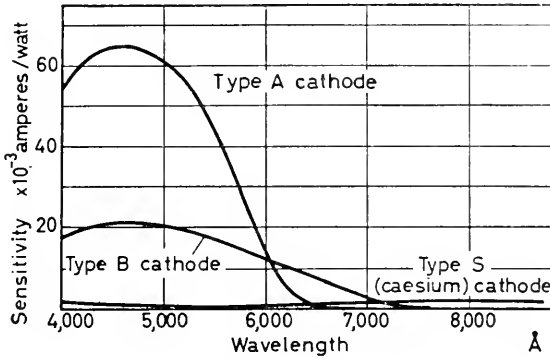


Figure 28.23 The spectral sensitivity curves for type A, B and S photocathodes

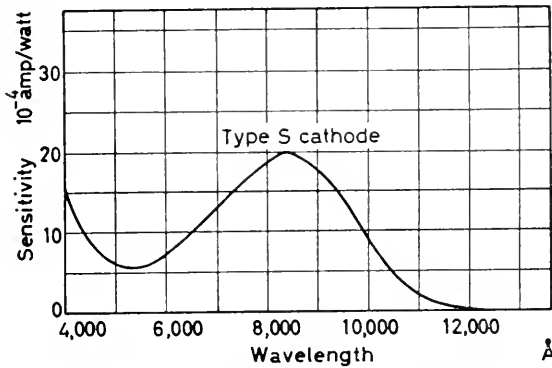


Figure 28.24 The spectral sensitivity curve of the type S cathode shown in Figure 28.23 redrawn on a larger scale

about 1,100 $m\mu$ where type A and B photocathodes are insensitive. In Figure 28.24 the spectral sensitivity of the type S cathode is redrawn on a larger scale.

In Figures 28.23 and 28.24 the photoelectric current produced at each wavelength has been calculated for an 'ideal' light source radiating equal amounts of energy at each wavelength; but in Figure 28.25 the relative sensitivities of the various photocathodes to a tungsten light source of C.T. 2,600°K are shown. As such a source radiates maximally at 1,000 $m\mu$ and little in the violet, the type S cathode is preferable to the types A and B, in spite of the much greater quantum efficiency of the latter. The ratio of the total areas covered by the curves in Figure 28.25 represents the ratio of the sensitivities of the respective photocathodes in mA/lm to a tungsten source. Although there are large individual variations of sensitivity among cells,

LIGHT SOURCES AND DETECTORS

even between those manufactured in the same batch, their sensitivities usually lie between the limits shown below:

<i>Type of cathode</i>	<i>Sensitivity in $\mu\text{A}/\text{lm}$</i>
Type A (S4)	30 to 100
Type B (S8)	10 to 60
Type S (S1)	10 to 60

One other photocathode worthy of mention is type S3. It is a silver, rubidium-oxide, rubidium cathode, with a maximum response near $420\text{ m}\mu$,

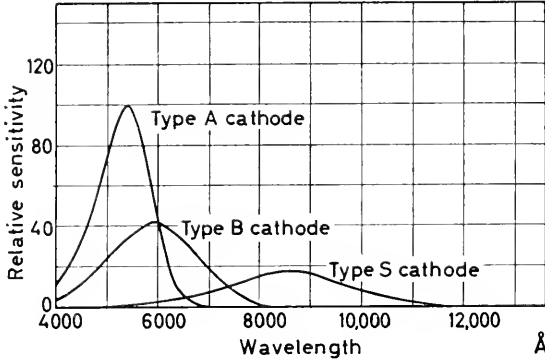


Figure 28.25 Comparison of the sensitivities of the three types of cathode to tungsten light of C.T. 2,600°K

in a lime glass envelope. It has a high sensitivity through the visible spectrum and it approximates to the spectral sensitivity of the eye. This photocathode is sometimes preferred for photometric measurements of atypical light sources, although it is more usual today to use type B cathodes with suitable correction filters. The filters are chosen such that the overall spectral response approximates to that of the eye.

Vacuum versus gas filled photocells

The simple photoemissive cell can be obtained with either a vacuum or a gas filling. It is important to study the different properties which these two types of filling impart to the cell in order to select the correct type for a particular application.

In some photocell applications a linear relationship is required between light intensity and photocell current over a wide range of illumination. In such cases a well-designed vacuum photocell is ideal. However, the detection of low light intensities with a vacuum cell is technically difficult due to the very small currents involved which require to be measured accurately. If a linear response is required at low illumination intensities the photomultiplier type cell, which is described later, is more satisfactory.

For the detection of moderately low light intensities the gas filled photocell has some advantages, although there is not a linear relationship between light intensity and photocell current. The greater sensitivity of the gas filled cell is due to the collision of the electrons emitted from the cathode with the gas molecules in the cell. This results in ionization of the gas and the

PHOTOELECTRIC DETECTORS

liberation of secondary electrons. Thus an amplification of the primary photocurrent occurs; the anode potential must exceed a critical value but not be so high as to produce irreversible total ionization of the gas. By gas filling it is possible to achieve an amplification of 10 or more over a reasonable range of light intensity. In this way sufficient current may be available to work relays, etc., with little additional thermionic valve amplification.

The sensitivity of a vacuum cell is independent of the anode potential, providing it is higher than the saturation potential of the cathode, so that close control of anode voltage is usually unnecessary; but the amplification factor of a gas filled cell does depend on the anode voltage and tight control is required if a constant sensitivity is desired.

Gas filled cells have a poor frequency response. Sensitivity falls off appreciably above 1,000 c/s. The frequency response of a vacuum cell can extend to many megacycles per sec if care is taken in the design to keep inter-electrode capacity low.

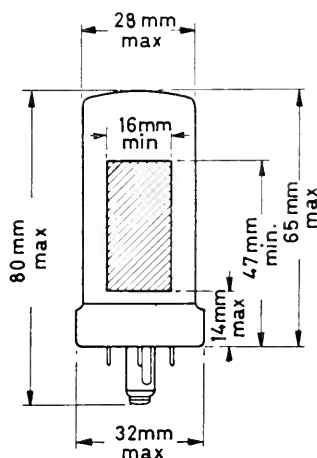


Figure 28.26 Dimensions of Mullard type 20CG and 20CV photoemissive cell

The number of different size photoemissive cells commercially available is very large indeed and reference to manufacturers' data sheets is required in order to select a suitable cell for a particular need. Two examples of a popular size photoemissive cell will be given as they illustrate the performance of this group of detectors.

The Mullard type 20CV is a vacuum photoemissive cell with the dimensions shown in *Figure 28.26* and type 20CG is a gas filled cell of equivalent dimensions. Both these cells have a type S1 photocathode so that they are particularly suitable for use with an incandescent light source or for the detection of radiation out to 1.1μ . No satisfactory photomultiplier cell has been marketed with this type of spectral response.

Figure 28.27 shows the performance of an average 20CV cell under a variety of illumination levels. It can be seen that, for a given light flux, saturation occurs with an anode voltage of the order of 25 V and that an increase in anode voltage above this level has little effect on the photocurrent.

LIGHT SOURCES AND DETECTORS

It follows that, providing the anode voltage is kept above this saturation voltage, the anode current will be proportional to the light flux falling upon the cathode. The voltage applied to the anode of a vacuum cell should not normally exceed 100 V nor should the cathode current exceed $20 \mu\text{A}$ for the type 20CV cell. The dark current for a cell of this type is less than $0.05 \mu\text{A}$.

Many users will not wish to detect or measure light intensity by measuring

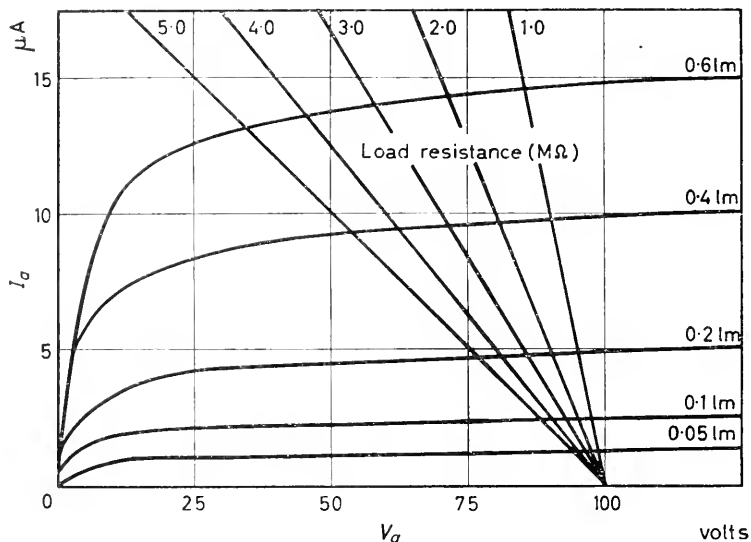


Figure 28.27 Performance data of the Mullard 20CV photocell. Anode current plotted against anode voltage with total illumination as parameter

current of the order of $10 \mu\text{A}$ and would prefer a voltage output. This is readily done by applying the anode voltage through a high resistance so that any increase in anode current will produce a voltage drop across the resistance. If a linear response is required the resistance should be chosen so that the anode voltage remains above the saturation voltage. In this way a voltage swing of some 75 may be achieved over the working illumination range. In practice it is difficult to use anode loads higher than 10 megohms. Reference to Figure 28.27 will assist in selecting a suitable anode load. The 20CV has a sensitivity of $25 \mu\text{A}/\text{lm}$ when the whole cathode (projected area 6.7 cm^2) is illuminated by a lamp of C.T. $2,700^\circ\text{K}$ and tested with a series resistance of 1 megohm.

The characteristics of the gas filled photocell 20CG is shown in Figure 28.28. Under favourable conditions the sensitivity of this cell is $150 \mu\text{A}/\text{lm}$ or about 5 times that of its vacuum counterpart. It should be noted that the maximum current ($5 \mu\text{A}$) which may be drawn from the cathode is considerably reduced in order to prevent damage to the cathode by ionization of the gas filling. The anode voltage must not exceed 90 V as complete ionization or 'gas glow' may occur with resulting damage to the cathode caused by the heavy current. This may occur even in the absence of light. The sensitivity of the cell varies with the anode voltage, being greatest at the point just

PHOTOELECTRIC DETECTORS

below that at which 'gas glow' occurs. It is therefore not possible to get a linear response to illumination in a cell with an anode load. The degree of gas amplification will also vary due to the absorption of the gas on to the electrodes and the glass envelope. This results in shorter life and unpredictable variations in sensitivity although these may be of small amplitude. A gas filled cell if continuously operated at its maximum rated voltage may decrease in sensitivity by 50 per cent in 500 hours of use. Much longer life

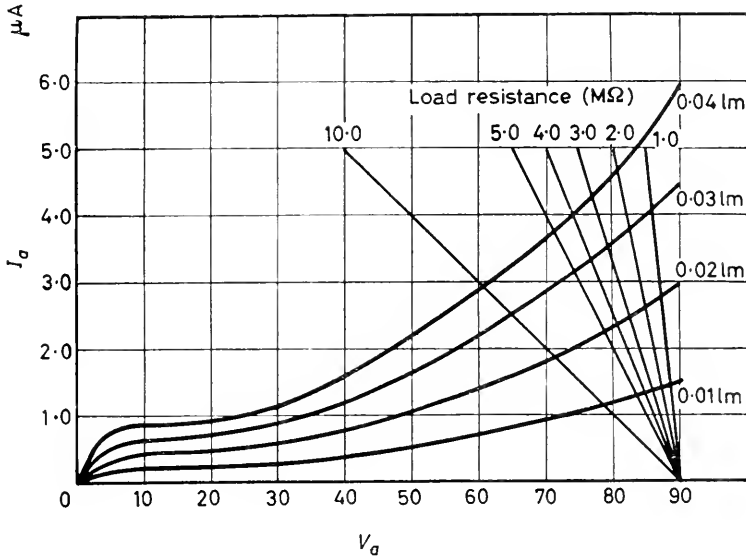


Figure 28.28 Performance data of the Mullard 20CG photoemissive cell. Anode current plotted against anode voltage with total illumination as parameter

can be obtained by reducing anode voltage or current, and by running the cell intermittently.

To summarize, it may be stated that the increased sensitivity of the gas filled photoemissive cell is obtained only by some sacrifice of linearity, maximum output current, stability, life and frequency response.

Electron-multiplier photoemissive cells

The simple types of photoemissive cell dealt with above usually require some degree of voltage or power amplification before their output signals can be examined or recorded. If the experiment only requires the observation of an alternating component in the output this can be readily achieved by means of a high gain a.c. amplifier whose input is capacitor-coupled to the photocell anode. A load resistor is, of course, connected in series with the cell. However, if d.c. components require amplification practical difficulties arise in constructing a sensitive d.c. amplifier which is free from 'drifts'. When dealing with low light intensity signals, whether alternating or not, the photomultiplier tube comes into its own, as the necessity for high gain amplifiers is avoided.

LIGHT SOURCES AND DETECTORS

The purpose of the electron-multiplier type of cell is to amplify the primary photocurrent within the cell itself. This is achieved by utilizing the emission of secondary electrons from another electrode (or several other electrodes in cascade) which is bombarded by primary electrons emitted from the photocathode. Tubes are available with as many as 14 secondary electrodes, usually called dynodes, and amplification factors as high as 10^8 can be readily achieved.

Probably the best-known photomultiplier is the R.C.A. 931A, which was extensively used during World War II and is still available from war surplus

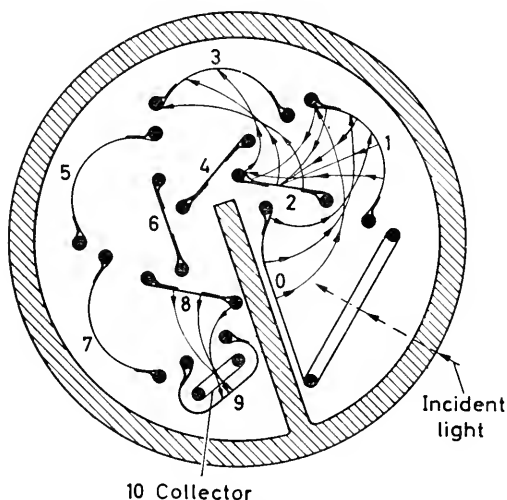


Figure 28.29 Electrode arrangement in R.C.A. type 931 photomultiplier
(By courtesy of Electronics¹⁰)

equipment. Figure 28.29 shows the general arrangement of the electrodes and electron pathways. Incident illumination passes through a wire grid to fall on the photocathode. The emitted photoelectrons are then guided electrostatically on to a secondary-emitting cathode, or dynode. An increased stream of electrons then passes to the next dynode and so on through the cascade to be collected at the ninth electrode, usually referred to as the anode. With a good 931A an amplification of 10^6 may be achieved with a voltage difference of 100 between adjacent dynodes. Thus a photosensitivity of 20 A/lm is obtained from a cell whose photocathode sensitivity is only 20 μ A/lm.

Two useful types of unit for supplying a photomultiplier are shown in Figure 28.30. If about 1,000 V is applied to the ends of the resistor chain, each dynode will receive about 100 V. It is wise to limit the anode voltage to about 50 V as this diminishes the danger of overheating the anode if an excessive light input is accidentally applied. The average anode current should not exceed 1 milliamp.

The logarithm of the amplification factor varies inversely with the log of the applied volts between adjacent dynodes (or 'volts per stage'), as shown in

PHOTOELECTRIC DETECTORS

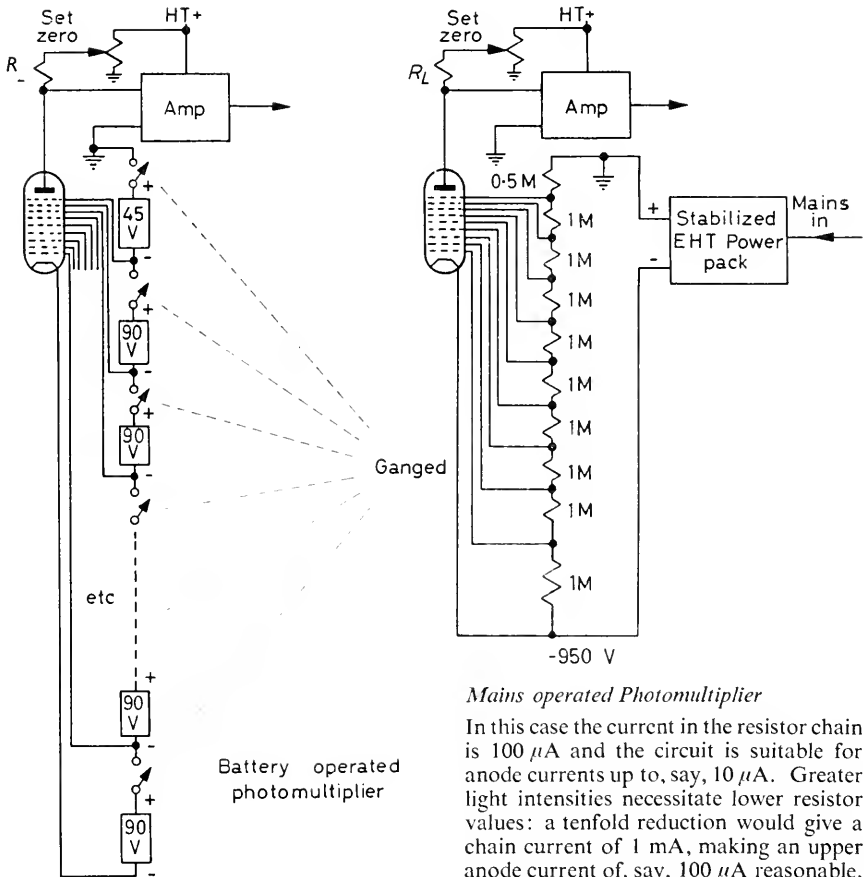


Figure 28.30 Supply circuits for securing linear amplification from a photomultiplier cell. The sensitivity is proportional to the anode load R_L , whose upper limit is set by the upper limit of frequency response required. If C is the amplifier input and stray capacitance, including the capacitance of any cable joining P.E. cell to amplifier, and if the upper turnover frequency is to be ω_{\max} , then $R_{L\max} = 1/(C\omega_{\max})$. Practical R_L 's range from 10 k Ω to 1 M Ω

Figure 28.31. If a constant sensitivity is required, care must therefore be taken to supply the multiplier tube with a constant voltage. Conversely this relationship between voltage supply and amplification can be used to advantage for the design of a wide range photometer as shown in the circuit of Figure 28.32. Here the supply voltage is altered to give a constant anode current for varying light intensities. The supply voltage required to give this constant current can then be measured; it will bear a logarithmic relationship to the light intensity.

If a multiplier cell is placed in total darkness a residual current will be found to flow. This is usually called the 'dark current', and is of the order of 0.1 μA in a 931A at 100 V/stage, although big variations between individual tubes occur. Some of this dark current is due to leakage at the base of

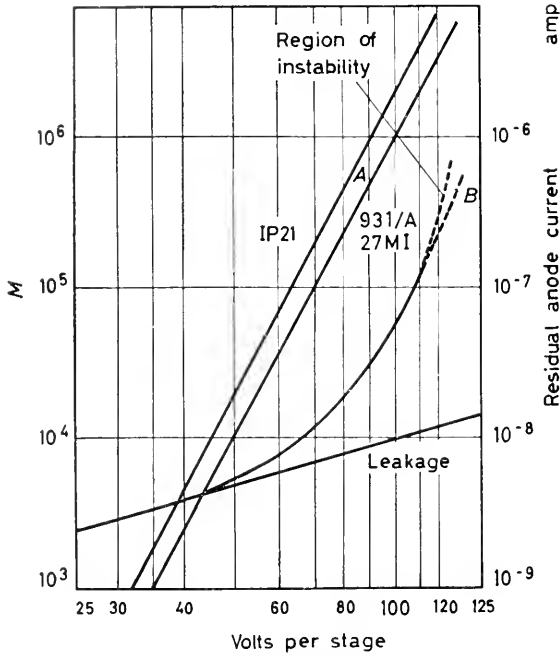


Figure 28.31 Typical amplification characteristics of 931A and IP21 photo-multipliers. A average multiplication; B typical residual current values as functions of stage voltage

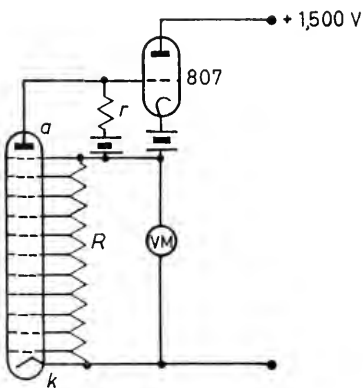


Figure 28.32 Basic circuit of a feedback densitometer (after Sweet¹¹)

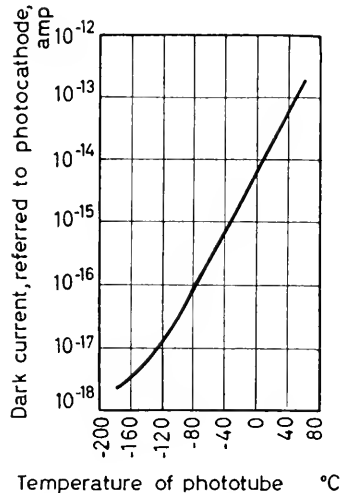


Figure 28.33 Effective cathode dark current of a multiplier phototube with antimony-caesium photocathode as function of temperature. (By courtesy of the Journal of the Optical Society of America¹²)

PHOTOELECTRIC DETECTORS

the tube (unfortunately the anode pin is adjacent to the cathode pin). This leakage can sometimes be reduced by cutting a groove in the base between the pins and coating with a silicone preparation. Most of the dark current is due to thermionic emission from the photocathode and dynodes. Reference to *Figure 28.31* will show that the dark current increases rapidly as the interstage voltage approaches 110. At higher voltages the photocell will become very 'noisy' and may show ionization. The maximum signal-to-noise ratio is obtained at about 90 V per stage. Some trial and error experiments are required for each individual tube to ascertain the optimum conditions.

The thermionic emission from the photocathode depends upon its temperature and a considerable improvement in performance can be achieved by cooling the cell to -78°C with solid carbon dioxide or -198°C with liquid nitrogen (*Figure 28.33*¹²).

To detect very low-intensity light signals it is of course an advantage to chop the light beam at some fixed frequency and then pass the a.v. signal from the photocell through a high- Q amplifier tuned to that frequency. If dark current noise is displayed on a C.R.O. it appears as 'grass'; each 'blade' differing in size from the next in a random fashion. Noise due to an electron being liberated from one of the higher dynodes is of smaller amplitude than an electron liberated from the photocathode. By counting the impulses greater than a certain amplitude appearing over a period of time, first with the cell in darkness and then with the light signal on the photocathode, and then comparing the counts, very great sensitivity can be achieved. This counting technique yields very high signal to noise ratios, since the effective frequency band-width is equal to the reciprocal of the counting time.

The output of a photomultiplier is linearly related to the light intensity over a wide range provided care is taken to ensure that the dynode voltages do not change with dynode current. The current flowing in the bleeder chain shown in *Figure 28.30* should be at least ten times greater than the maximum current drawn by the last dynode. At high light intensities non-linearity will occur due to space charge effects in the last stages.

The frequency response of these tubes is very wide and can extend to several megacycles; the transit time of the electrons in the tube limits the response.

Fatigue effects in the 931A are small providing photocurrents not exceeding $10\ \mu\text{A}$ are drawn and the overall voltage is kept down to 700: the fatigue tends to occur during the first 30 minutes but variations are found from tube to tube.

The IP21 is a specially selected cell of the 931A type and should be used for the detection of signals of low intensity. Mullard 27M2 is similar to the 931A. The 27MI is a specially selected and tested version of the 27M2 and has a signal-to-noise ratio about 10 times better. These four tubes have a type S4 spectral sensitivity. IP22 has a type S8 sensitivity. IP28 and 27M3 have an ultraviolet transmitting glass envelope which enables signals with wavelengths down to $220\ \text{m}\mu$ to be detected. Worthy of special mention are the E.M.I. group of photomultipliers. These tubes use a semi-transparent end window photocathode which absorbs some of the radiant energy striking it. Photocathodes with areas ranging from 78 to $9,700\ \text{mm}^2$ are manufactured. The dynodes have a 'Venetian blind' type of structure consisting of

LIGHT SOURCES AND DETECTORS

TABLE 5

Tube type	No. of stages	Cathode diameter mm	Envelope diameter max mm	Length sealed ± 3 mm mm	Min. overall sensitivity 160V per stage or less A/L	Typical overall sensitivity at 160V per stage A/L	Minimum photosensitivity μ A/L	Typical photosensitivity μ A/L	Dark current at minimum overall sensitivity μ A	Window glass	Basic cathode type
6099B	11	111	130	230*	100	400	25	40	0.5	Pyrex	Cs-Sb
6099F	11	111	130	230*	25	—	35	45	0.5	Pyrex	Cs-Sb
6099C	11	111	130	230*	5	—	10	—	0.2	Pyrex	Cs-Sb
6097B	11	44	51	112	200	600	30	50	0.1	Pyrex	Cs-Sb
6097S	11	44	51	112	200	400	20	25	0.1	Pyrex	Cs-Sb(S)
6097F	11	44	51	112	50	—	40	55	0.1	Pyrex	Cs-Sb
6097C	11	44	51	112	10	—	10	—	0.04	Pyrex	Cs-Sb
9514B	13	44	51	121	2,000	6,000	30	50	1.0	Pyrex	Cs-Sb
9514S	13	44	51	121	2,000	4,000	20	25	1.0	Pyrex	Cs-Sb(S)
9514C	13	44	51	121	100	—	10	—	0.4	Pyrex	Cs-Sb
6094B	11	10	50	94	200	600	30	50	0.01	Kodial	Cs-Sb
6094C	11	10	50	94	10	—	10	—	0.004	Kodial	Cs-Sb
9502B	13	10	50	108	2,000	6,000	10	50	0.1	Kodial	Cs-Sb
9502C	13	10	50	108	100	—	10	—	0.04	Kodial	Cs-Sb
6255B	13	44	51.5	121*	2,000	6,000	30	50	1.0	Quartz	Cs-Sb
6256B	13	10	51.5	108*	2,000	6,000	30	50	0.1	Quartz	Cs-Sb
6095B	11	44	51	112	200	300	20	30	0.1	Pyrex	Bi-Ag-Cs
6095F	11	44	51	112	50	—	30	—	0.1	Pyrex	Bi-Ag-Cs
6095C	11	44	51	112	10	—	10	—	0.04	Pyrex	Bi-Ag-Cs
9524B	11	21	28.5(1 $\frac{1}{8}$ "	112	100	—	20	—	0.05	Lime Soda	Cs-Sb(S)
†6260B	11	44	50	112	200	500	20	25	0.1	Kodial	Cs-Sb(S)
†6262B	14	44	50	125	10,000	30,000	20	25	5.0	Kodial	Cs-Sb(S)

* ± 6 mm.

† Normally supplied for replacement purposes only

PHOTOELECTRIC DETECTORS

oblique slats which permit the secondary electrons to pass from one dynode to the next (*Figure 28.34*). Tubes with 11, 13 or 14 dynodes are available. Gain varies approximately as the eighth power of the overall voltage supply for an 11 stage tube and a highly stabilized supply is necessary for consistent results. A 14 stage tube with a 160 V per stage may have a sensitivity of between 10,000 and 30,000 A/lm with a dark current of about 5 μ A.

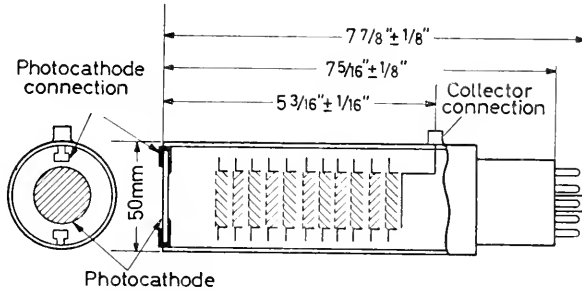


Figure 28.34 Arrangement of electrodes of E.M.I. type 5311 photomultiplier

Table 5 gives the main characteristics of this range of cells. Their spectral sensitivity differs slightly from the sensitivity curves shown in *Figures 28.23* and *28.24* because of the form of photocathode used. Refer to Rodda¹³ and Sommer³ for design theory for these tubes.

Photomultiplier tubes should not be operated in the presence of strong magnetic fields unless protected by a highly permeable—e.g. mumetal—shield. Any metal clamps around the body of the tube should be at cathode potential to prevent small electrical leaks which may give rise to light emission and to prevent deflection of the electron beams within.

In selecting a photocell for a particular task it is important to bear in mind that the thermionic noise from the photocathode is proportional to its area. If it is possible to focus the light signal optically by means of lenses or mirrors so that it forms a small image, a photocathode about the size of the image should be used as this will give the best signal-to-noise ratio. If it is not possible to focus the light on to a small area, as in, for example, a scintillation counter, then a large area photocathode will be preferable.

Selenium barrier-layer cell

This type of cell has the great advantage over other light detectors that it does not require an external source of voltage; the action of the absorbed radiation is to generate an e.m.f. within it. It is the cell used in the familiar photographic exposure meter.

Its structure is shown in *Figure 28.35*. The cell consists of an iron or steel base plate upon which is deposited a specially prepared selenium compound in a crystalline state. On top of this is deposited a thin transparent, electrically conducting layer. This layer and the base plate are the two electrodes. To facilitate electrical contact with the upper layer a robust metal ring is sprayed on to it. The front surface, excluding the collecting ring, is then covered with a protective lacquer whilst the exposed iron is protected by a rustproof metal coating. Potted cells, cast in Araldite resin behind a glass

LIGHT SOURCES AND DETECTORS

window and enclosed in a metal or bakelite housing, are also manufactured (Megatron Ltd.), which enable them to be used in the presence of moisture or corrosive vapours.

When light falls on the cell the upper surface becomes electrically negative with respect to the base plate. If an external load is connected across the electrodes a current will flow; thus the cell may be connected directly to a suitable galvanometer or meter and the resulting current measured.

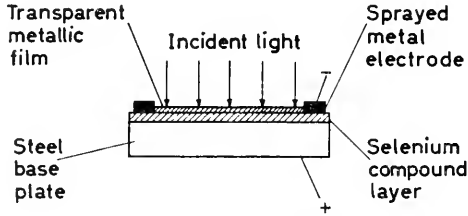


Figure 28.35 Cross-section through a selenium barrier-layer photocell

A convenient equivalent circuit is shown in Figure 28.36. A current generator i_p is shunted by a capacitor C and an internal resistance R_i in series with a resistance R_s . If the external load resistance is R , the current through it becomes:

$$i = \frac{i_p R_i}{R_i + R_s + R}$$

R_s is mainly the value of the thin collecting electrode and is about 50 ohms in most cells. C varies from 0.1 to 0.5 μF ; depending upon the cell

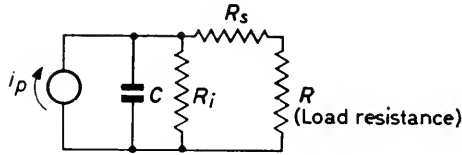


Figure 28.36 Equivalent circuit of a barrier-layer cell

area. This capacity limits the frequency response of the cell. i_p and R_i vary with the intensity of the incident illumination. The relationship between illumination and R_i for a 45 mm diameter Weston cell is shown in Figure 28.37. It can be seen that the equivalent internal resistance R_i varies both with the external load R and with the illumination. The internal resistance rises to a high value in darkness and this permits a number of cells to be connected in parallel without mutual short circuiting when they are independently illuminated.

The maximum potential generated with high illumination is about 0.4 V and it varies non-linearly with light intensity. The potential does not however vary much between cells of different area. The low voltages and relatively low impedances associated with barrier-layer cells render them unsuitable for use with valve amplifiers. If further power amplification is required transistors may conveniently be used (e.g. page 697).

PHOTOELECTRIC DETECTORS

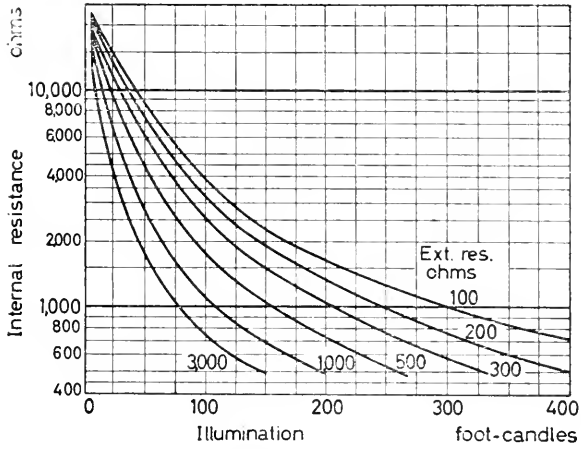


Figure 28.37 Relationship between illumination and change of internal resistance for a 45 mm diameter Weston barrier-layer cell

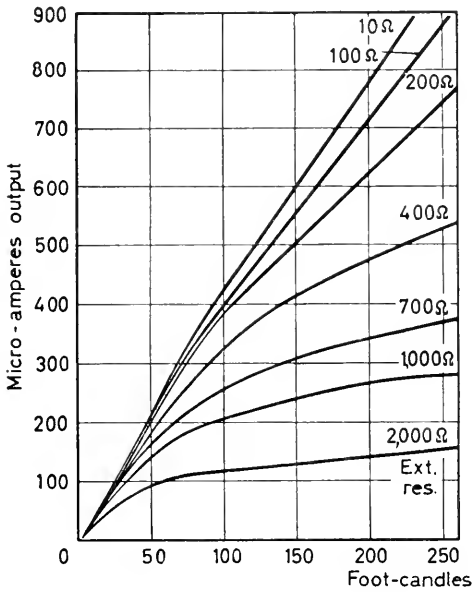


Figure 28.38 Relationship between illumination and current output for a 45 mm Weston barrier-layer cell

LIGHT SOURCES AND DETECTORS

The relationship between current flow and illumination for different values of external resistance is shown in *Figure 28.38*. It can be seen that if a low resistance measuring instrument is used a fairly linear relationship exists between current and illumination intensity (Weston, 45 mm diameter cell). Individual cells, even from the same manufacturer, vary widely, however, in output. Short-circuit current sensitivities vary from 100 to 600 $\mu\text{A}/\text{lm}$.

The selenium barrier-layer cell generates a remarkable amount of power. A Weston 45 mm diameter cell with an external load of 350 ohms and an illumination level of 200 foot-candles would deliver 100 μ -watts to the load (*Figure 28.39*). Their efficiency as energy converters when working under optimum conditions is about 2.5 per cent, therefore their quantum efficiency must be quite high.

By selecting an appropriate load the effect of ambient temperature changes can be minimized as demonstrated in *Figure 28.40*.

These cells suffer from fatigue effects. At worst, the output may increase or decrease by about 4 per cent depending as usual upon the individual cell. When high accuracy is required, the current output at a given illumination should be plotted against time for an individual cell and the time to reach equilibrium noted. Further readings should only be taken after this initial time has elapsed. The amount of fatigue which occurs depends not only upon the intensity of the light, but also upon the colour, being more marked with light from the red end of the spectrum. The long-term stability of barrier-layer cells is good even after many years of continuous use provided they are protected from very high illumination levels and from heat. Use of a heat filter is sometimes advisable if long exposure to tungsten light sources is required.

Due to the high internal capacity of these cells their frequency response is poor. At 3,000 c/s the response is down by about 5 per cent and at 10,000 c/s it is down by approximately 80 per cent. Small area cells have a better frequency response than large.

Typical spectral sensitivity response curves for a number of selenium barrier-layer cells are shown in *Figure 28.41*. The response in the near ultraviolet region is fairly good and is sometimes useful. For example, the output tends to be high when a cell is exposed to violet-rich skylight out of doors, when compared with tungsten illumination of equal intensity indoors. This is convenient for a photographer using panchromatic film, which has a similar violet sensitivity; but it is inconvenient when measuring the equivalent *visual* intensity of the light. If an Ilford filter type 827 is placed before it, its response curve approximates to that of the eye. To obtain an even closer fit, more complex filters are required (Preston¹⁴). Megatron Ltd. manufacture a selenium cell which is sensitive out to about 900 $\text{m}\mu$ in the near infra-red. It is a cell of high sensitivity, especially when used with a tungsten light source.

The selenium cell is particularly valuable as a portable photometer when used in conjunction with a suitable meter. It may be conveniently used in the laboratory for approximate photometry as its spectral response curve is near to that of the eye, especially if suitable colour filters are added. It has also been used as a portable generator of electrical energy to supply power to

PHOTOELECTRIC DETECTORS

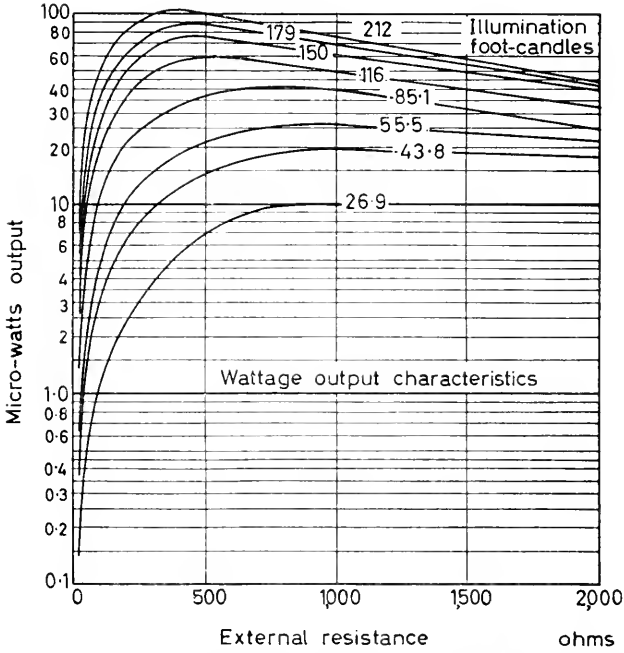


Figure 28.39 The energy output of a 45 mm diameter Weston cell for various illumination levels and load resistances

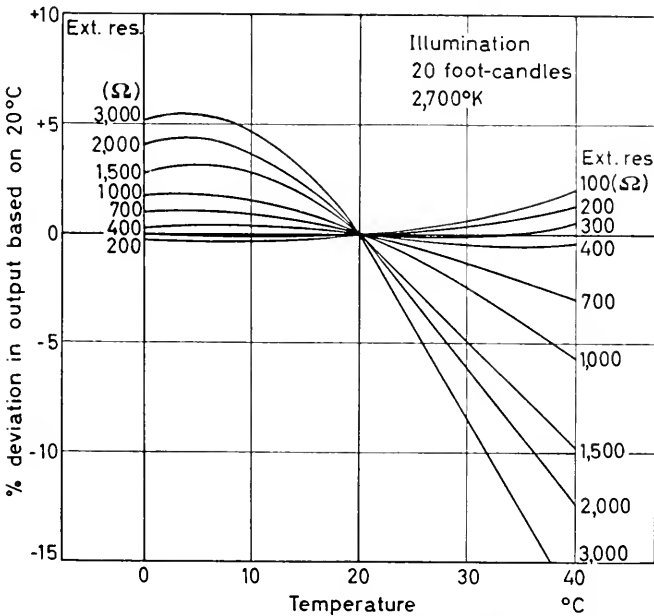


Figure 28.40 Effect of temperature on current output at 20 foot-candles for a 45 mm diam. barrier-layer cell

transistor circuits. The power available from the sun's radiation at ground level is about 1 kW/m^2 and the popular name for the barrier-layer cell is becoming 'the sun battery'.

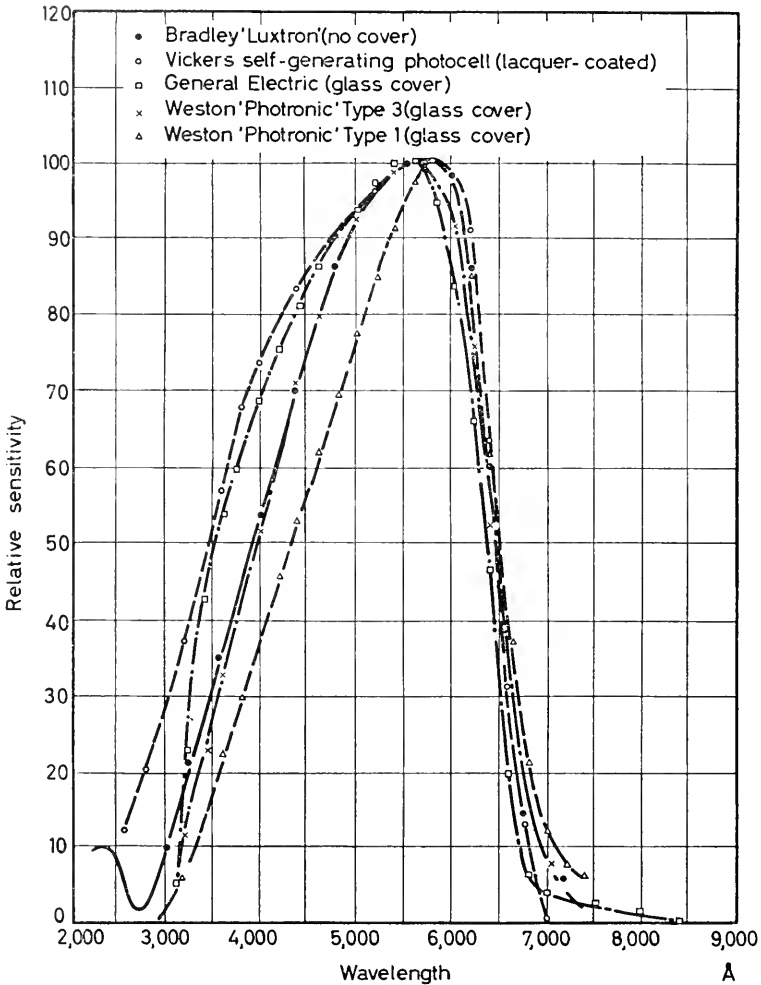


Figure 28.41 Spectral response of various selenium barrier-layer cells of American manufacture

Photoconductive cells

Several substances exhibit the property of altering their electrical resistance when radiant energy falls upon them, and this enables them to be used as photodetectors. Interest in the detection of infra-red radiation during World War II resulted in the development of a number of photocells in this class. As light from a tungsten source contains much of its radiant energy in the infra-red, these detectors work well in conjunction with a tungsten lamp. The lead sulphide cell is the best known.

PHOTOELECTRIC DETECTORS

Lead-sulphide photoconductive cell

The lead-sulphide cell consists of a layer of specially prepared lead sulphide a few microns thick deposited on glass and sealed in a protective mount.

The Mullard 61SV is a typical example of an uncooled lead cell having a high sensitivity at room temperatures. The sensitive area is 6 by 6 mm, and the device has a resistance of about 6 megohms in darkness. The spectral response curve is shown in *Figure 28.42*. It can be seen to have a peak at 2.5μ with a range from 0.3 to 3.5μ .

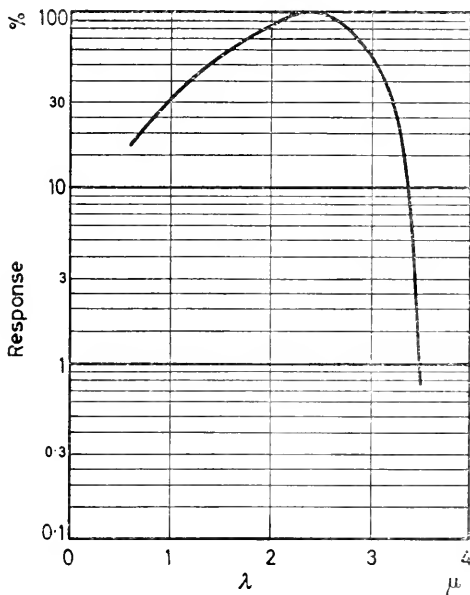


Figure 28.42a Spectral response curve of the Mullard 61SV lead-sulphide photoconductive cell

The cell may be used connected to a direct voltage source and the change in resistance detected by placing a meter in series; or a more satisfactory arrangement might be to place the cell in one arm of a Wheatstone bridge and then back off the standing current which would otherwise flow through the meter when the cell was in darkness. A chopped light source may be used and the alternating portion of the signal may then be fed by means of a coupling capacitor to an a.c. amplifier.

In *Figure 28.42* is shown the output of a 61SV under various loads, voltages, and light intensities. If tested with a lamp of C.T. $2,700^{\circ}\text{K}$, the sensitivity is found to be about 3 ma/lm. Due to the good infra-red performance of these cells, a source at a temperature of only 100 to 200°C can be detected. Without any optical system such a cell could detect the presence of a 350°C soldering-iron at a distance of 100 yards. The signal conversion factor with respect to 200°C is shown in *Figure 28.43*.

If the 61SV is cooled, an increase in signal-to-noise ratio can be obtained, as shown in *Figure 28.43*. The British Thomson-Houston Co., Ltd. make a

LIGHT SOURCES AND DETECTORS

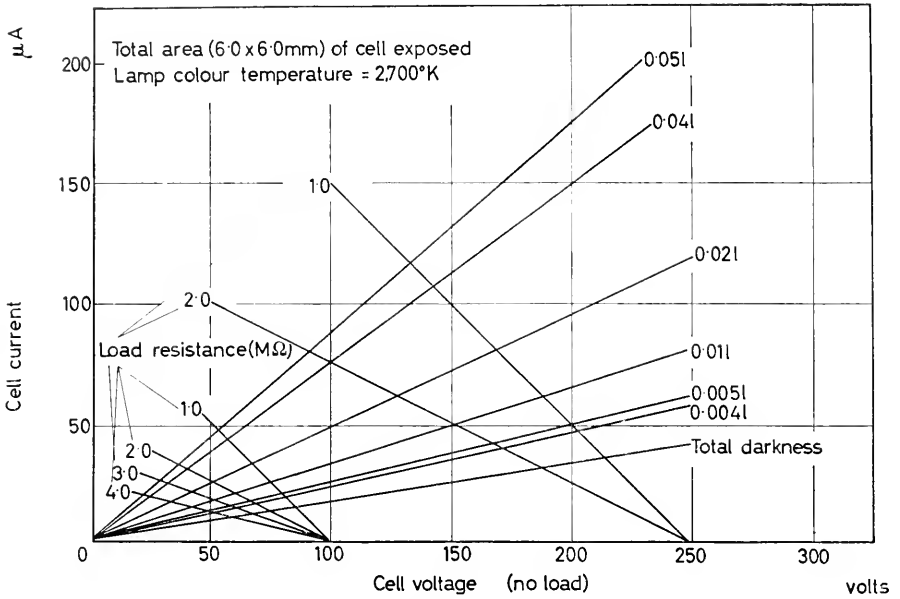


Figure 28.42b Output data of a Mullard 61SV lead-sulphide cell. Cell current plotted against no load cell voltage

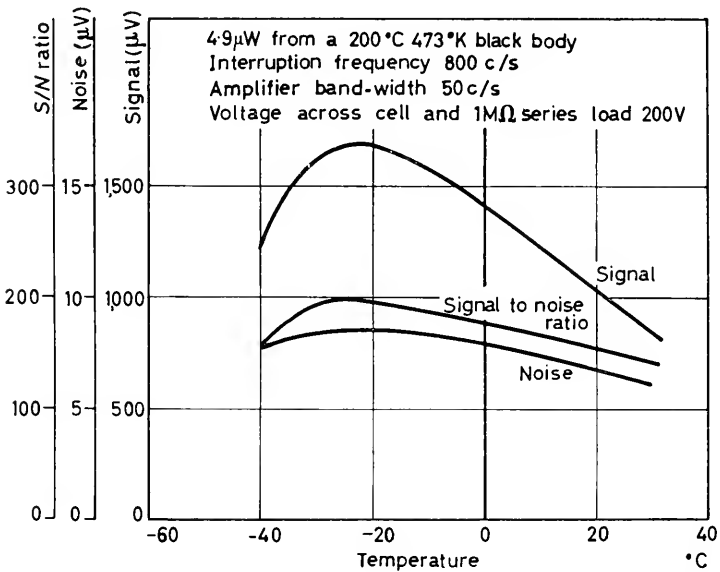


Figure 28.43a Signal conversion factor plotted against black body temperature or a 61SV Mullard lead-sulphide cell

PHOTOELECTRIC DETECTORS

lead-sulphide cell which is enclosed in a transparent Dewar flask which may be filled with solid carbon dioxide to reduce the temperature to -78°C .

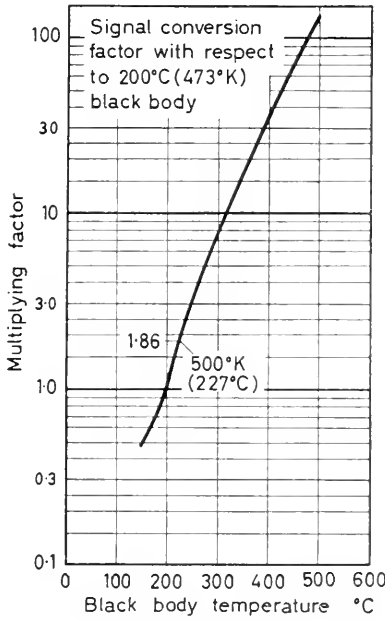


Figure 28.43b Effect of temperature on the performance of a Mullard 61SV lead-sulphide cell

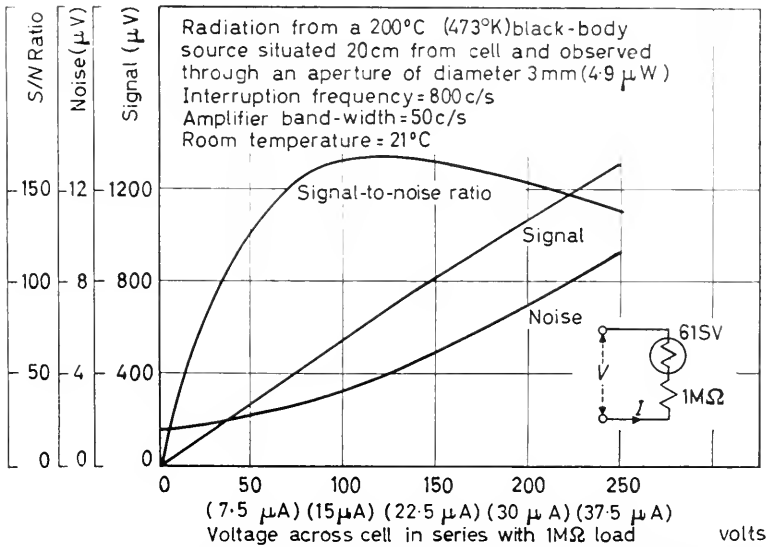


Figure 28.44 Effect of applied voltage on performance of Mullard 61SV lead-sulphide cell

The variation of signal, noise and signal-to-noise ratio with applied voltage is shown in Figure 28.44.

LIGHT SOURCES AND DETECTORS

To detect radiant energy beyond 3μ a Mullard 61RV lead-selenide uncooled cell may be used (*Figure 28.45*); alternatively, the 63TV lead-

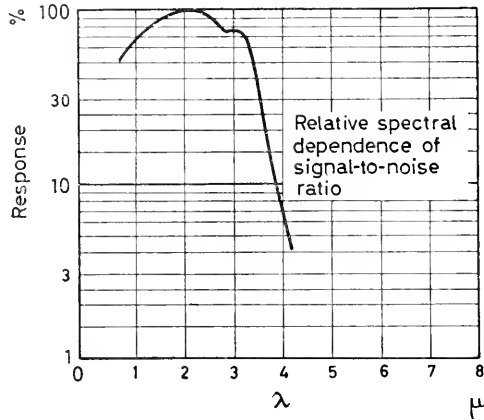


Figure 28.45 Spectral sensitivity curve of a 61RV Mullard lead-selenide cell

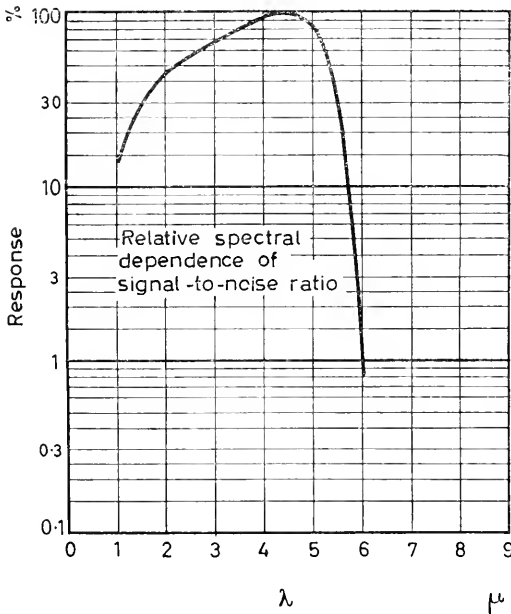


Figure 28.46 Spectral response curve of a Mullard 63TV cooled lead-telluride cell

telluride cell is available, which requires liquid O_2 or N_2 for its operation (*Figure 28.46*). The 61RV has a time constant of $15 \mu\text{sec}$, and the 63TV, 10 to $100 \mu\text{sec}$.

PHOTOELECTRIC DETECTORS

Area n - p junctions can also be formed in lead sulphide by ensuring a sulphur deficient surface on a p -type single crystal. These photodiodes have spectral characteristics similar to those for the photoconductive type and have the great advantage of a time constant of about $1 \mu\text{sec}$: the sensitivity is correspondingly lower. These cells are of the photovoltaic type and operate without an external power source.

Cadmium-selenide cell

A useful photoconductive cell with a high internal amplification has been developed by E. Schwartz of Messrs Hilger and Watts, Ltd. It consists of a

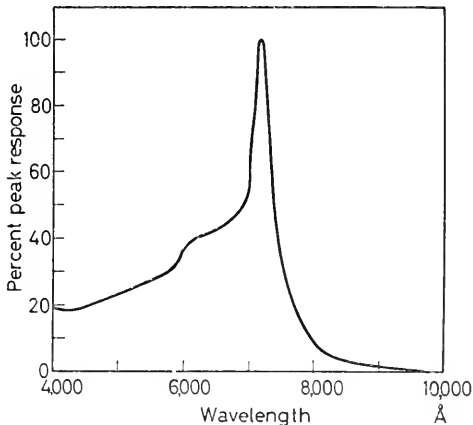


Figure 28.47 Spectral sensitivity of the Schwartz cadmium-selenide cell (Hilger and Watts, Ltd.)

polycrystalline layer of treated cadmium-selenide on an insulating base, the whole being placed in a small protective mount. The most sensitive cells have a photosensitive surface of width only 0.2 mm and a length of 1.5 or 10 mm , and it is usually necessary to focus the light on to this narrow area. Double cells in a single mount are also produced for comparison work or for use as movement detectors.

The spectral sensitivity of the Schwartz cell is shown in *Figure 28.47*. The sensitivity extends below $400 \text{ m}\mu$, through the ultraviolet, down to X-ray wavelengths. There is a sharp peak at $710 \text{ m}\mu$, with a fairly rapid cut-off thereafter.

As these cells are of the photoconductive type they require a source of power to activate them; however, only low voltages are required. The characteristics of a cell with a sensitive area of 1 by 0.2 mm are shown in *Figure 28.48* for a variety of illumination intensities. It can be seen that for a high illumination of, say, $1,000 \text{ lux}$ ($= 200 \mu\text{lm}$ on the cell) about 1 mA will flow with a voltage of 4 . This is equivalent to a sensitivity of 5 A/lm , with a dark current of $0.6 \mu\text{A}$. At the low illumination level of 1 lux ($= 0.2 \mu\text{lm}$ on the cell) about $50 \mu\text{A}$ will flow with an applied voltage of 30 . This gives a sensitivity of 250 A/lm with a dark current of $5 \mu\text{A}$. These cells have therefore a very high internal amplification and they can be used for energizing relays, etc., directly. If the incident energy cannot be focused as a small

image on the cell, then cells with a larger photosensitive area are available. For example the 10×10 mm cell will pass 10 mA in an illumination level of 100 lx, with a dark current of $6 \mu\text{A}$ when working on 6 V, giving a sensitivity of 1 A/lm.

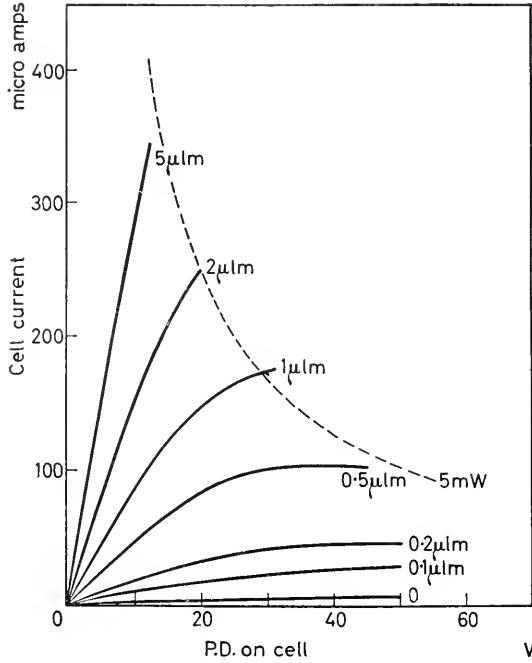


Figure 28.48 Characteristic curves for a Schwartz cadmium-selenide cell

The time constant of Schwartz cells varies with the size of the sensitive area and is of the order of 1 to 10 msec. No definite information is available on fatiguing tendencies but it may be significant that the manufacturers advise that these cells are intended for monitoring purposes only, and not for the quantitative measurement of light intensities.

The General Electric Company Ltd. are developing cadmium-sulphide photoconductive cells of two types—a miniature single-crystal cell and a large-area powder-layer cell. Various spectral sensitivities can be obtained by varying the production technique. Peaks lying between 500 and 700 μm have been produced. The sensitivity to tungsten light is about 1 A/lm. Mullard are developing similar powder-layer cells which at low illumination may be used to operate relays directly.

Germanium diffusion-junction photocells

Standard Telephones and Cables, Ltd., are now manufacturing germanium diffusion-junction photodetectors which have the great advantage of small size and large current or voltage output. The P50A has an average effective sensitivity area of 1.5 by 0.1 mm. The unit is enclosed in a small sealed tube 5.6 mm in diameter and 8.8 mm long.

PHOTOELECTRIC DETECTORS

The spectral response of the P50A is shown in *Figure 28.49*. It can be seen that the response fits well the spectral output of a tungsten source, so that good transfer of energy from source to detector may be expected.

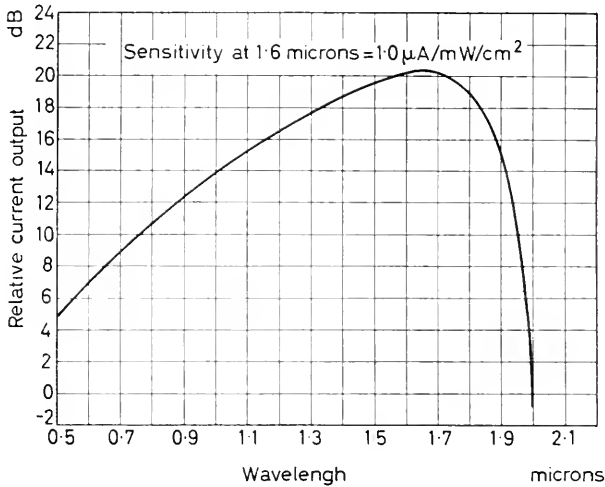


Figure 28.49 Spectral response curve of a typical P50A Standard Telephones and Cables germanium diffusion-junction photodetector

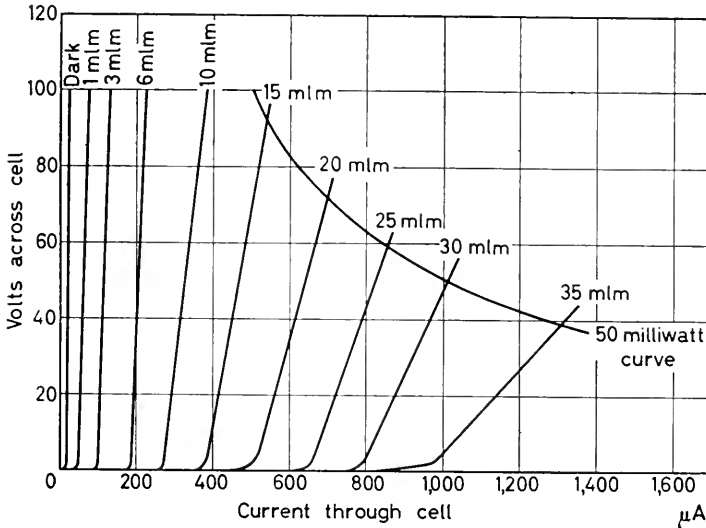


Figure 28.50 Performance data of an average P50A germanium phototransistor

The cell must be used with an electrical power source of the correct polarity. *Figure 28.50* shows the current versus applied voltage characteristics for an average cell at various light levels at a temperature of 20°C . At least 1 V must be applied before the cell will function. The effect of a 100 kilohm load is also shown. Due to the small area of the active photosensitive element (0.15 mm^2) the sensitivity of the device to unfocused illumination is

LIGHT SOURCES AND DETECTORS

poor compared with large-area photodetectors such as the selenium photo-emissive cell. However, a good signal-to-noise ratio can be obtained with focused light beams. The noise produced by the cell is little greater than that to be expected from Johnson noise alone. The sensitivity of the cell to oblique illumination is shown in *Figure 28.51*.

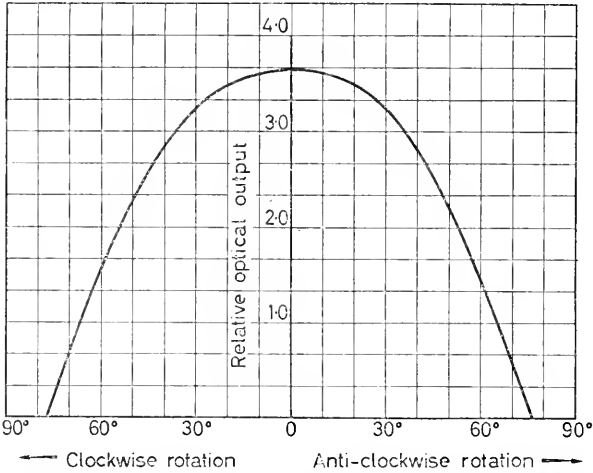


Figure 28.51 Relative output of P50A at various angles of light incidence

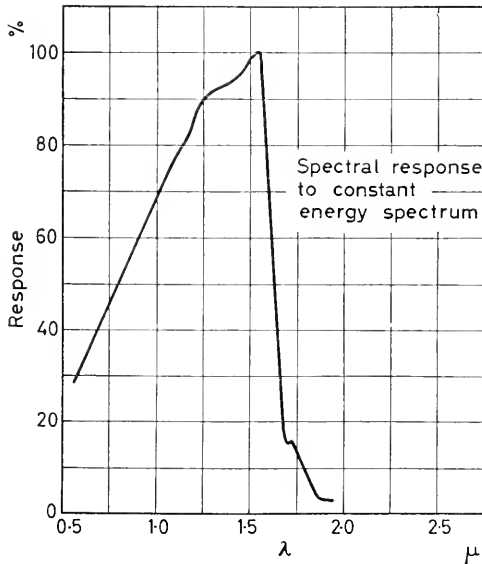


Figure 28.52 Spectral response curve of the Mullard type OCP71 phototransistor

One of the main disadvantages of this type of detector is the variation of the characteristics with ambient temperature. This should be borne in mind when considering any application.

PHOTOELECTRIC DETECTORS

These cells are clearly of value in 'on or off' relay applications where the light flux may be focused on the cell. Their small size makes them particularly suitable for use in portable equipment, although their temperature characteristics may be a disadvantage. Their excellent high-frequency

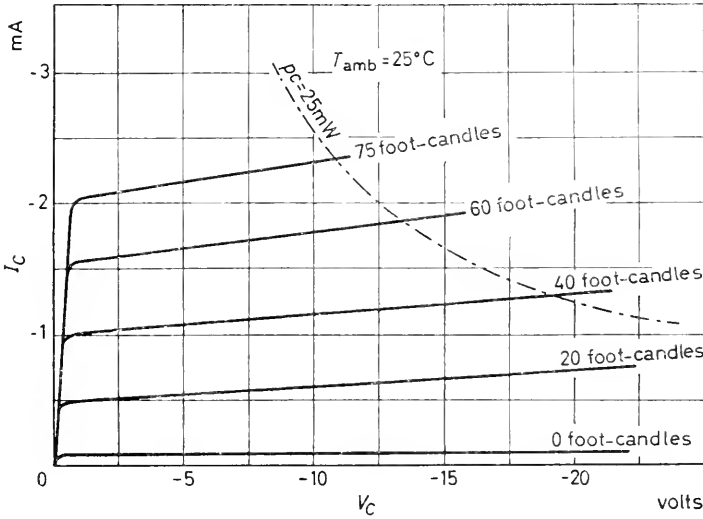


Figure 28.53 Operating characteristics of the Mullard OCP71 phototransistor

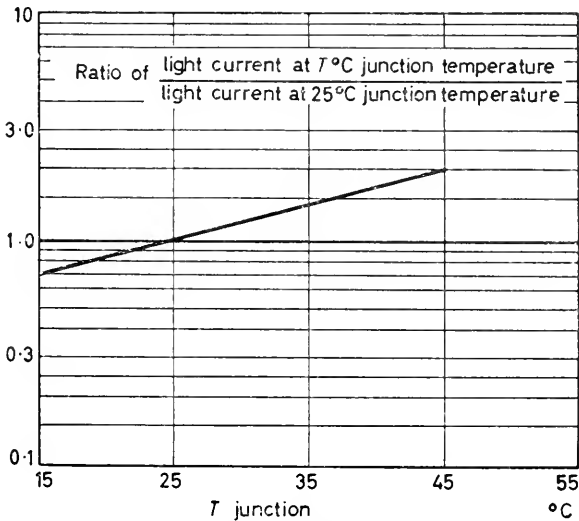


Figure 28.54 The effect of ambient temperature in the output of the Mullard OCP71 phototransistor

response is valuable in modulated light systems, or for high-speed counting, especially where an infra-red sensitivity is required.

The OCP71, made by Mullard Ltd., is a phototransistor of the *p-n-p* alloy

type, hermetically sealed in an all-glass envelope of frosted glass. Dimensions are: diameter, 5.9 mm, length 15 mm. The photosensitive region is 7 mm² in area. Due to the diffusing properties of the glass envelope the cell is sensitive to radiation reaching it either end-on or from directions around the envelope. The spectral response is shown in *Figure 28.52*. The cell must be used with an electrical power source of correct polarity, connected to the emitter and collector leads. *Figure 28.53* shows the current versus applied voltage characteristics for an average OCP71 at various levels of illumination. As can be seen from *Figure 28.54* the characteristics of the OCP71 vary markedly with temperature.

REFERENCES AND BIBLIOGRAPHY

- ¹ WALSH, J. W. T. *Photometry* London; Constable. 1953
- ² SPENCER, D. A. *Colour Photography in Practice*, 3rd ed. London; Pitman. 1948
- ³ SOMMER, A. New multiplier phototubes of high sensitivity *J. sci. Instrum.* 27 (1950) 113
- ⁴ BOURNE, H. K. *Discharge Lamps* London; Chapman & Hall. 1948
- ⁵ MARTIN, L. C. *Technical Optics*, Vols 1 and 2. London; Pitman. 1953
- ⁶ BEESON, E. J. G. A xenon lamp for colour photography *J. Photogr. Sci.*, 1, No. 3 (1953)
- ⁷ BALLANTYNE, D. W. G. The phenomenon of electroluminescence *Marconi Review* 19 (1956) 160
- ⁸ BOWTELL, J. N. Electroluminescence and its applications *J. Instn. elect. Engrs.* 3 (1957) 454
- ⁹ KAZAN, B. and NICOLL, F. H. *J. Opt. Sci. Amer.* 47 (1957) 887
- ¹⁰ RAJCHMAN, J. A. and SNYDER, R. L. *Electronics* 13 (1940) 20
- ¹¹ SWEET, M. H. *J. Soc. Mot. Pict. and Tel. Engrs.* 54 (1950) 34
- ¹² ENGSTROM, R. W. Multiplier phototube characteristics *J. opt. Soc. Amer.* 37 (1947) 420
- ¹³ RODDA, S. *Photoelectric Multipliers* London; Macdonald. 1953
- ¹⁴ PRESTON, J. S. The relative spectral response of the selenium rectifier photocell in relation to photometry and the design of spectral correction filters *J. sci. Instrum.* 27 (1950) 135
 FELLGETT, P. A current stabilized photomultiplier power supply *J. sci. Instrum.* 31 (1954) 217
 HOLIDAY, E. R. and WILD, W. An image converter to extend the useful range of photomultipliers to longer wavelengths. *J. sci. Instrum.*, 28 (1951) 282
 LE GRAND, Y. *Light, Colour and Vision*. London; Chapman & Hall. 1957
 NÁRAY, Z. On the reduction of the dark current of photomultipliers *J. sci. Instrum.* 33 (1956) 476
 PRESTON, J. S. The specification of a spectral correction filter for photometry with emission photocells *J. sci. Instrum.* 23 (1946) 211
 SUMMER, W. *Photosensors* London; Chapman and Hall. 1957
 WILSON, R. Constancy of photomultiplier gain *J. sci. Instrum.* 30 (1953) 472
 WRIGHT, W. D. *Photometry and the Eye* London; Hatton Press. 1949
 ZWORZYKIN, V. K. and RAMBERG, E. G. *Photoelectricity* New York; Wiley. 1949

MEASUREMENT AND CONTROL OF TEMPERATURE

J. W. L. BEAMENT

INTRODUCTION

The measurement and control of temperature is a practical problem in almost all biological work, although usually an adjunct to a particular investigation. Electrical methods of control are virtually the only ones in use—the most refined methods involve electronic control gear. For many purposes, electrical methods involving electronics offer very great advantages in temperature measurement too, and for very accurate purposes are again essential. This chapter starts with a basic summary of the essentials of heat, temperature, etc., which is followed by a discussion of methods of measuring temperature, together with practical details of the equipment; the third section discusses the theory of methods of temperature control, and is followed by examples of equipment layout for the principal biological needs—the temperature controlled environment of air or of water.

TEMPERATURE AND HEAT

Energy transformation

For all practical purposes, a current I passing through a resistance R produces a quantity of heat, I^2R ; current and resistance are readily controlled quantities and since this heat is produced without gaseous by-products, etc., electrical heating is the only one envisaged in biological work except for such crude purposes as the heating of laboratories and greenhouses.

Temperature

Temperature is the thermal equivalent of electrical potential; when two bodies of different temperature are in contact, heat tends to flow from the higher to the lower temperature; generally speaking the rate of transfer is proportional to the temperature difference, regardless of size and the 'amount' of heat in the two. Without recourse to very complex methods, such as are used in refrigeration processes, *only* a temperature difference can promote such a passage of heat—a body can otherwise only be heated by causing it to transduce another form of energy put into it into heat.

Thermometry

Many physical properties change in a reliable, repeatable and reversible way with temperature changes; those which are amenable to accurate mensuration are thus useful as indicators of temperature. A *thermometer* displays the change produced by temperature change; as we are not interested

in the temperature of the thermometer as such, but of the object in contact with the thermometer, a thermometer must equilibrate to the temperature of its surroundings rapidly—which usually means it must be a good conductor of heat—and in that process it must not change the heat content of its surrounds appreciably or the resultant temperature measured will not be that prevailing before the thermometer was introduced.

Temperature scales

It is known that water in contact with melting ice, under stated conditions, has a constant and repeatedly obtainable temperature. The thermal exchange of equilibrating a thermometer with an ice-water mixture involves only a change in the amount of ice—not of the temperature. Hence this is used for reference purposes as the 'lower fixed point' of temperature. Similarly, a 'higher fixed point' is obtained from water boiling at atmospheric pressure. On the Centigrade scale, these are the 0° and 100° points; on the Fahrenheit scale they are called 32° and 212° . Once these two points have been calibrated on the instrument whose temperature-dependent property is being used as a thermometer, the intervening change of property is divided into *equally spaced* intervals: on the Centigrade scale into 100 parts, on the Fahrenheit into 180. No account is taken of whether the addition of equal quantities of heat produces successively similar changes of property—for example, the mark 50°C on the mercury-in-glass thermometer is half way between the 0 and 100° marks, on the stem; 50°C on the platinum resistance thermometer is half way in resistance between that at 0 and 100°C . However unless the laws governing the change of volume, and of resistance, with successively added quantities of heat are identical, 50°C will not be the same temperature on both scales—and in fact, readings taken in the same medium around 50°C do vary between the two, by almost one degree. Whenever specifying temperature to an accuracy greater than 0.5°C it is necessary to state the measuring element used, and this is especially important if the measurement was made on an instrument other than a mercury thermometer, or one not calibrated against a mercury thermometer.

Heat quantity

The thermal capacity of a body is the quantity of heat needed to raise its temperature through one degree. For unit mass of material the figure is called the specific heat of the material: thus the product of specific heat and mass gives the thermal capacity. Specific heats vary widely—that of water is highest, while metals are commonly low. With a heterogeneous collection of objects, such as might constitute a biological environment, either in water, or more particularly air, it must be appreciated that in raising the temperature of the whole collection the added heat energy has to be distributed very unevenly according to the specific heats of the objects.

Conduction and insulation; convection and radiation

The thermal conductivity of a material gives the rate at which heat flows through it for a unit temperature gradient across it. Generally speaking, good thermal conductors are also good electrical conductors, while electrical insulators are also poor heat conductors, and thus difficult to heat,

MEASUREMENT OF TEMPERATURE

but useful as insulators. Many heat 'insulators' do not however rely entirely on the thermal conductivity of their material, but on cellular air spaces and a restriction of convection. When fluid is heated locally it expands in volume; the attendant decrease in density causes it to tend to rise in the surrounding cooler material. Circulation of thermal origin set up in this way is called convection: such convection speeds the heating of a body of fluid appreciably, for clearly it takes place faster than the heat energy can be transmitted through the fluid by conduction. Gases, such as air, are extremely poor conductors of heat, but this property can be utilized for insulation purposes only when the air is sealed in small cells to prevent convection—hence the great value of expanded plastic foams, cork, slag wool, etc. Conversely, while convection may assist in the heating of a fluid environment it can only take place where temperature differences exist, thus indicating a lack of temperature uniformity. Thermal gradients set up by convection can amount to several degrees Centigrade change. Heat can also be transferred in the form of electromagnetic radiation—typically the infra-red. The important point to realize is that many materials, including air, transmit heat in this form without being appreciably heated themselves; hence if a visibly red heater is placed in a box of air, an appreciable portion of the heat emission is absorbed directly by the walls of the box without heating the air. On the other hand, light (as well as mechanical energy) entering an environment is normally converted into heat, so that radiation may prove a source of unwanted heat in a controlled enclosure.

MEASUREMENT OF TEMPERATURE

Temperature is measured by a thermometer, which to most people is a mercury-in-glass thermometer. The universal use of this machine and its 'convenience' is accepted. The outline of limitations which may lead to incorrect information from the mercury thermometer, and which may make it necessary to use one of the admittedly less convenient electrical methods, is given below.

Limitations of the mercury-in-glass thermometer

(a) Contact; while most biological work will involve the measurement of the temperature of fluids, anyone who has tried to measure the temperature of a solid will know the inconvenience of the device.

(b) Thermal capacity; for any reasonable volume change of mercury the bulb must be large—hence the thermal capacity is correspondingly large; thus equilibration will be slow, and according to the size and nature of the body may lead to an appreciable heat exchange and an alteration of the body's temperature. This is particularly true of a body such as air. The most sensitive type of thermometer—the Beckmann—is capable of readings more accurate than 0.01°C but it measures only temperature differences; reference calibration is very difficult. The mercury thermometer is useless for following rapid temperature changes.

(c) Hysteresis; the glass envelope of a thermometer has appreciable hysteresis if subjected to a rapid temperature cycle; this can materially affect measurements of accuracy greater than 0.1°C : permanent distortion is unlikely.

MEASUREMENT AND CONTROL OF TEMPERATURE

(d) Orientation; sensitive mercury thermometers will only give reproducible readings in one (usually the vertical) position.

(e) Remote reading; the instrument is not readily adapted to give the temperature of sealed boxes, especially light-tight containers.

(f) Recording; the instrument does not lend itself to automatic continuous recording.

(g) Most delicate measuring instruments are themselves delicate; glass, and the hazard of metallic mercury released from a broken thermometer are particular problems.

(h) The mercury-in-glass system is regularly used as a thermostat element; it is not easy for the amateur to construct such thermostats.

Other non-electrical thermometers

Unfortunately for biologists, almost all the research and development in temperature measurement and control has been carried out industrially for temperatures very much higher than any he is likely to encounter. Of industrial types, the only ones he might find used are the bimetal, in which the differential expansion of two metals is magnified to display temperature, and the gas and vapour-pressure thermometers, which use the expansion of vapour sealed in a bellows as the sensitive element. These devices are of no use in sensitive work; they have the advantage only of robustness, large energy of movement and economy. Their mechanical movement is often directly linked with a recording pen in thermographs, where they have application in crude meteorology.

Electrical methods

The thermocouple—When two wires of different materials are joined together at either end, giving two junctions, and these are maintained at different temperatures, an e.m.f. is produced which causes a current to flow round the loop (*Figure 29.1*). It is considered that two processes

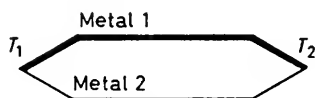


Figure 29.1

contribute to this potential: the Peltier effect, due to the junctions, and the Thomson effect, due to the temperature differences between the ends of the individual wires. The resultant potential, the sum of these two effects, is approximately linearly related to temperature difference, for the predominant Peltier effect is directly proportional to temperature difference, while the Thomson effect is proportional to the difference in the squares of the temperatures. The universality of the device is realized when it is known that thermocouples are regularly used to measure temperatures up to 1,600°C in industry; yet they are equally one of the most accurate ways of measuring temperatures in the biological range, and they can follow small changes rapidly and accurately.

The advantages of the thermocouple are: the junctions can be minute,

MEASUREMENT OF TEMPERATURE

they can be inserted into small spaces and parts of small animals and plants; their thermal capacity is also very small and they can therefore equilibrate rapidly without altering the temperature of small objects.

A variety of metals can be used, and some typically useful ones (e.g. platinum-rhodium systems) are biologically inert. With certain precautions, they are remote reading instruments. They are without hysteresis, or other aging affects.

The disadvantages are: Materials such as platinum, rhodium, are usually expensive; for sensitive work, the voltages to be measured accurately are minute, and a high-grade galvanometer is needed with attendant delicacy and expense; a cold-junction reference temperature must be maintained with an accuracy of an order better than the measurement being taken; elaborate compensation precautions may be involved in the circuitry, or it may be necessary to thermostat the galvanometer and remaining accessory equipment.

Practical considerations of the thermocouple—A closed loop of wire containing a circulating current is not very much use; yet as soon as any measuring device is placed in the circuit, containing, as it will, a number of metals, such as copper, brass, constantin, etc., in series, we have a loop containing many junctions (*Figure 29.2*). If each of these junctions contributed their own Peltier effects, the original e.m.f. across the junctions

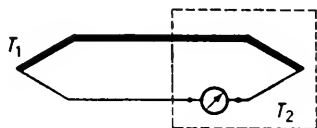


Figure 29.2

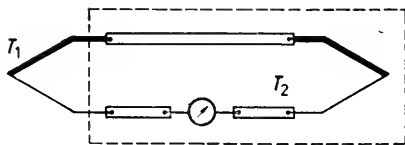


Figure 29.3

could never be realized. The law of intermediate metals states that provided the temperature remains unchanged any different metal, or series of metals, may be placed in one part of the leads forming a pair of junctions without changing the overall e.m.f. It is not only important to keep the temperature of the measuring instrument constant for this reason—the voltages generated by thermocouples are of the order of microvolts per degree—so that one has to use a really sensitive instrument which will only give stable readings if its temperature is kept reasonably constant; all its component parts and the leads in the circuit will have temperature-sensitive resistances. It is thus desirable to use massive copper leads, keeping down the additional resistance of the circuit and also saving long leads of expensive materials such as platinum. The meter should be protected by a shorting switch which is opened to take readings. *Figure 29.3* indicates the parts of the circuit which should not undergo much temperature change. In addition, leads and contacts must be kept thoroughly clean (e.g. copper contacts should be scraped), and the humidity should not be too great, to prevent surface films of water and voltaic cells at metal junctions.

The typical order of voltage to be measured is $50 \mu\text{V}$ per degree C difference in junction temperatures. At first sight it seems that if this is displayed on a $50 \mu\text{V}$ meter, a reading to 0.01°C is quite straightforward, and a reading

to 0.001°C is not over difficult. However, while such temperature *differences* can be measured, having the two junctions in two respective places, the experimental problem becomes very much different as soon as one wants to calibrate the device, or if one wishes to measure one temperature absolutely to any degree of accuracy; for the only reference temperatures one can obtain very accurately are those of melting ice and boiling water. Now thermocouples have to be calibrated against something—for accurate work a Beckmann thermometer is usual. The problem is that it is difficult to obtain any metering device of which the scale can be read more accurately than to 1 part in 1,000; this is one degree C full-scale deflection for a sensitivity of 0.001° , while if one wished to read temperatures around 20°C and had to display ice temperature to 20°C on the scale, one would not be able to read more accurately than to 0.02°C . The sensitivity will correspondingly go down as the temperature range is increased. If in fact this is still of sufficient sensitivity, then a calibration curve is constructed by setting a Beckmann thermometer in an ice-water mixture containing both junctions, while the meter or galvanometer is set to zero; the thermometer with one couple wound round the bulb is then placed in successively warmer water—giving the mercury plenty of time to equilibrate, and stirring the water well.

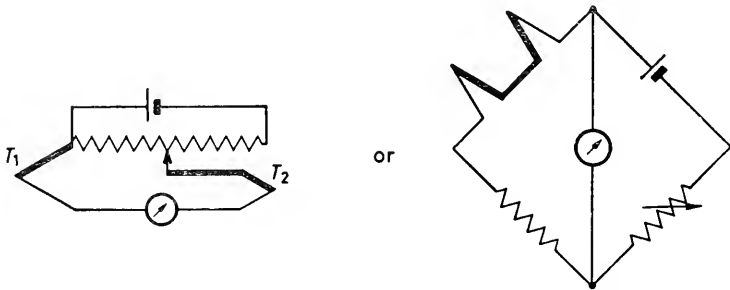


Figure 29.4 Basic null-point and bridge circuits for thermocouple measurements

If, however, it is necessary to work at higher sensitivity, then use can be made of the law of intermediate temperatures, which states that if a thermocouple develops a given voltage between temperatures T_1 and T_3 , this will equal the sum of the voltages developed between temperatures T_1 and T_2 , and T_2 and T_3 , where T_2 is any intermediate temperature. It is thus possible to calibrate a couple between 0°C and a suitable temperature above it, then to place the cold junction, together with a second reference thermometer, into high temperature water and continue the calibration curve in the form of additional voltages for known temperature differences between the mercury thermometers.

Alternatively, one can use a potentiometer system to 'back-off' the excess voltage due to the temperature difference (Figure 29.4). The couple can again be calibrated in steps, and for very sensitive work it is advisable to use a high resistance potentiometer to keep down the battery drain, and to include a standard Weston cell, in a separate circuit, with which to check the voltage of the ordinary cell (a standard cell will give a voltage of very

MEASUREMENT OF TEMPERATURE

great accuracy provided that it is limited to an infinitesimally small current drain).

Conclusions—The great advantage of the thermocouple is clearly its small size and thermal capacity, and its remote-reading characteristic; to use it as a device for the very accurate measurement of temperature is also possible, but a considerable technical feat in itself. Where a sensitivity of 0.1°C is all that is required, a couple of copper and constantan, soft-soldered or brazed at the junctions, is convenient. It is not normal to use thermocouples in combination with amplifiers, or for recording or controlling purposes, because of the low voltages and low impedances of the system (this is not, of course, true of the output of industrial instruments measuring temperatures of the order of $1,000^{\circ}\text{C}$ which have substantial outputs).

The thermopile—The output of a thermocouple can of course be increased additively, by placing a large number of them in series. For accurate work, the large number of leads between individual hot and cold junctions would make the device cumbersome, and of greatly increased thermal capacity. However, it has applications in detecting temperature differences at small distance separations, e.g. radiation receptors, where radiation falls on one set of junctions as a black body and the other set are shielded, or in some forms of hygrometer, where one set measures wet bulb temperatures and the other dry.

Resistance thermometers—The essential of the resistance thermometer is that it makes use of the temperature coefficient of resistance of a conducting element. Leaving the special case of the thermistor (see below) till later, the materials used as sensitive elements are almost exclusively the metals nickel, copper or platinum. Over the temperature ranges which will interest biologists the resistance of a metal changes linearly with temperature on the Absolute temperature scale to an accuracy of about 4 per cent (though of course on the platinum-resistance scale a platinum thermometer is self-linear).

The advantages of the resistance thermometer are: (a) remote reading; (b) it can be made of inert material, e.g. platinum; (c) it is very stable and shows almost no hysteresis with large rapid temperature excursions; (d) its minute thermal capacity is particularly suitable for measuring gas temperatures; (e) once calibrated, its reading is absolute (*cf.* a thermocouple's cold junction).

Its disadvantages are: (a) the resistance of the element may be changed by adsorbed contaminants; (b) it may not be used in contact with any conducting liquid such as water; (c) a very thin wire element is delicate; (d) if heated by the measuring current, it may give incorrect readings; (e) calibration is difficult.

Practical consideration—The accurate measurement of resistance is a straightforward physical problem, and for the resistance thermometer a Wheatstone bridge is advised (*Figure 29.5*). All leads must be of constantan or very heavy copper, to keep their own temperature-resistance changes minimal, and the comparator resistance must also be temperature stable. In order to neglect any electrochemical voltages or polarization in the circuit an a.c. bridge is recommended, and thus the comparator resistances must be non-inductively wound. It has already been pointed out that the

MEASUREMENT AND CONTROL OF TEMPERATURE

measurement current through the thermometer wire will heat it unless it is very small, the temperature measured will in fact be that of the heated wire less the cooling effect of its surroundings, and it may well heat its surroundings appreciably. This current must be kept to an absolute minimum.

A typical thermometer has a resistance of the order $100\ \Omega$, with an order of $0.4\ \Omega/\text{degree C}$ change. Readings to $0.01\ \Omega$ (0.025°C) should be quite straightforward, with the precaution that a voltage of $1\ \text{V}$ across the element will cause it to dissipate $10\ \text{mW}$ —a fine coil could easily rise 1°C when passing this current—and the actual temperature in the element will be different according to the thermal capacity and rate of fluid flow past it.

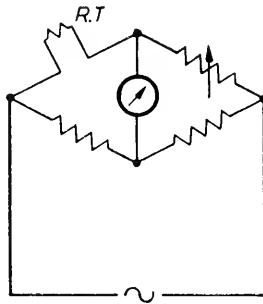


Figure 29.5 A.c. bridge circuit of a resistance thermometer (RT)

The resistance in series with the element must be at least of $1\ \text{k}\Omega$, or larger, in order to limit the current in the element to a negligible heat effect (i.e. less than a mW). The support for the very fine wire must also be minimal in size—mica sheet is suggested. Calibration against a mercury thermometer in a gas needs an extremely long time for the mercury to equilibrate; it is possible to calibrate in a bath of pure mineral oil provided: (1) it is ascertained that immersion has no effect on the resistance of the element and (2) that it is cleaned with chloroform and pure alcohol afterwards.

Conclusion—The resistance thermometer is an excellent accurate method for measuring gas temperatures and for following quite rapid fluctuations of temperature in them. The application of a self-balancing bridge would enable it to be used with recording gear, though it is rarely used as a control element.

The thermistor—This is a resistance thermometer of a very different kind which merits special treatment. It consists of a minute bead of heavy-metal oxide which has a large negative temperature coefficient of resistance and offers an extremely convenient method of measuring temperature to quite fine limits. It has a small thermal capacity and can be obtained sealed in a very thin glass capsule so that, unlike the platinum thermometer, it can be immersed in most liquids: it is therefore not subject to contamination effects, it is reasonably stable against aging and is certainly no less robust than a mercury thermometer or other glass instrument probe. A whole range of types and resistances is available ($2\ \Omega$ to $100\ \text{k}\Omega$ at 20°C). The resistance change is logarithmic and it approximately halves its resistance for every 20° rise in temperature; a combination of very high rate of

MEASUREMENT OF TEMPERATURE

resistance change and high impedance means that it can be used with long leads without problems of their actual resistance or temperature coefficients. Like other resistance thermometers, measurement involves passing a current through them, with the result that they dissipate power and heat themselves. Their power-dissipation must be limited to something of the order $1/10$ mW, and because of the great change in resistance over a reasonable temperature excursion it is wise to feed them from a constant current source. It must be appreciated that unlike the platinum thermometer this instrument can be calibrated in water, but whereas such a high specific heat liquid will keep the self-generated temperature down, air will not do so; hence unless the thermistor emits very little heat indeed its calibration will change with its surroundings. The thermistor has also some hysteresis if subjected to rapid large temperature fluctuations, but these will not seriously affect normal biological work: in general the instrument should not be subjected to a greater temperature gradient than 10°C per minute, nor to a greater overall excursion than $50\text{--}60^{\circ}$ without re-calibration. The shift in calibration due to more violent experience may be 1 per cent; under the best conditions measurement to 0.01°C may be obtained. Thermistors of any one type and resistance show individual differences of such an order that each must be separately calibrated against a mercury-in-glass scale.

Practical details—A thermistor is capacitive and should not be used in an a.c. bridge, for the capacity will also have a temperature coefficient; while a d.c. bridge technique might give the most sensitive results, methods in which the temperature is displayed on a meter will commend themselves for normal work. Two systems have been found reliable in practice, depending on the application of the machine: a battery-operated unit for portable field work, where the thermistor is used as an ecological probe; and a mains-operated unit, particularly suitable for bench work, for the measurement of air temperatures, and those of small animals and plants.

Mains unit—The device of Beament and Machin¹ has given satisfactory service. Here, the thermistor, fed from a constant current source to limit its power dissipation, controls the grid of a cathode follower. A second cathode follower (second half of a double triode) gives selected voltages, and these are compared by a microammeter between the cathodes. The meter is 'protected' against reverse voltages by a rectifier in parallel, though this must be chosen with care for it is temperature sensitive and its shunt value will vary. High cathode resistances together with very high resistances in the thermistor and comparator potentiometers reduce the effect of mains voltage fluctuations. The thermistor chosen for this unit is of the highest resistance range— $100\text{ k}\Omega$ at 20°C . The thermistor is calibrated against a good mercury thermometer in water, and the comparator resistances may have to be padded slightly in order to get overlap on the temperature scales. In the design shown (*Figure 29.6*) the machine switches 'on' into the highest temperature range, since the rectifier will always protect against low-temperature overload but not against high-temperature overload. A small amount of ripple in the power supplies does not seem to affect the instrument adversely.

Portable unit—This employs an unbalanced bridge technique; using the lowest range thermistor ($2\text{ k}\Omega$ at 20°C), the current drain is so small that

MEASUREMENT AND CONTROL OF TEMPERATURE

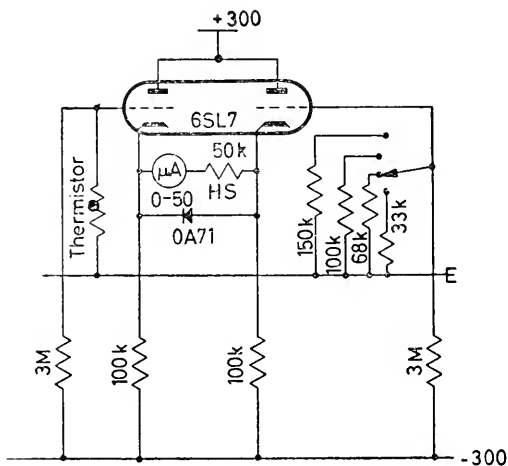


Figure 29.6 Practical circuit for measurement of temperature by a thermistor—mains-operated unit. Thermistor type: 100 k Ω at 20°C; suitable meter, 50 μ A in series with 50 k Ω . The selector switch covers the ranges 15–20, 20–30, 30–50, and 50–100°C

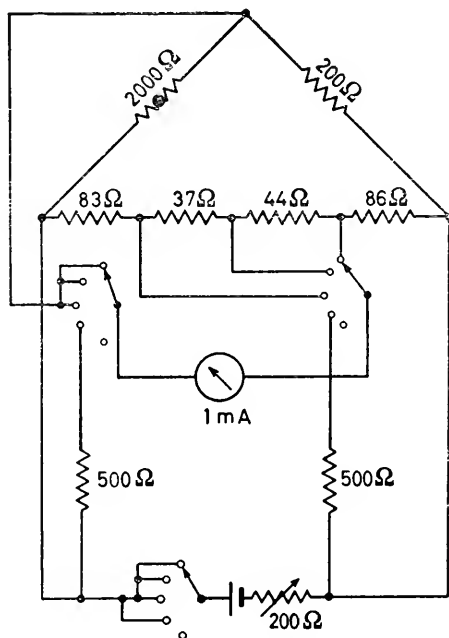


Figure 29.7 Practical circuit for measurement of temperature by a thermistor—portable battery-operated unit. Thermistor type: 2 k Ω at 20°C. In addition to three temperature ranges, a switch position is incorporated for setting the bridge voltage to a standard level by means of the variable resistance

TEMPERATURE CONTROL

a 1.5 V dry cell should operate for a few days and its change in voltage should not be more than the sensitivity of the measurement. In view of the isolated circumstances under which the machine may be used, a switching system uses the same meter to set up and check a reference battery voltage. Temperatures are displayed in three ranges, but by adjusting the values of the individual resistances (*Figure 29.7*), these can be set to requirement.

Recording of temperature

The common thermograph consists merely of a bimetal element directly driving through levers a recording pen on a clock drum. This is about the only thermal element with sufficient energy of change to drive a recorder direct; but the element has a very great thermal capacity and is only useful for meteorological work where changes are slow and accuracy does not need to be better than $\frac{1}{2}^{\circ}\text{C}$ at best. For more accurate work, or where rapid changes have to be followed, one is virtually limited to using an element giving rise to a proportional electrical output which can be linked to a pen-recorder (see transducers) or string-recorder. Both of these will need at least one d.c. amplifier stage beyond anything that a thermocouple or thermistor bridge can produce. A circuit described below under 'temperature control by variable heater' shows a thermistor controlling a two stage d.c. amplifier; its output (a power tetrode) could be connected to a standard recording device to provide a good temperature recording apparatus.

TEMPERATURE CONTROL

Introduction

To the biologist, temperature control will normally mean the maintaining of a fixed chosen temperature in an enclosure of water or air containing living material; this enclosure may vary from a constant temperature room or incubator running continuously at one pre-determined temperature to a small experimental enclosure in which a wide range of temperatures is to be produced. As accessories, he may be concerned with ovens, hot plates and autoclaves.

Theory of temperature control

If we consider a discrete body at ambient temperature and introduce heat energy into it its temperature will rise proportionally with the amount of heat added. As the temperature of the body rises, and a temperature difference is set up between it and the ambient, the body will lose heat into the ambient at an ever increasing rate and eventually a dynamic equilibrium is set up; for any given rate of heat input, and of ambient temperature, there will be a body temperature which will remain constant so long as the ambient temperature also remains so, but it will depend on ambient temperature as much as on heat input. The heat put into the body may be considered in two parts—that needed to raise the temperature of the body, due to its thermal capacity, and that continuously provided to replace losses into the ambient. At equilibrium, the whole heat input equals the heat loss.

Now if a body could be thermally isolated, the problem of temperature control would be simply a matter of inserting a suitable amount of heat

energy; thermal insulation cannot be at all complete, however, and for all practical purposes heat loss will be appreciable. The primary problem is to compensate for changes in the amount lost because of variations in the ambient (which may be due to radiant and convective loss as well as purely conductive loss). It might be suggested that the control mechanism should therefore follow fluctuations in ambient temperature, but this is not the basis normally used.

The main technical problem of temperature control, then, is to regulate the amount of, and rate of supplying, heat energy. For this an electric resistive heater is used, fed from a fixed supply voltage. When operating continuously the maximum body temperature will be obtained, in relation to a given ambient temperature; for all lower body temperatures the heater must operate discontinuously, and control consists of mechanisms to open and close the heater circuit appropriately.

Fixed-heater switching thermostat

Information on the temperature of the body is obtained by a temperature sensitive element (T.S.E.)—this is not required to present temperature in the sense of a thermometer, an independent thermometer should always be used for that purpose. A T.S.E. linked to a switch, or other control gear for maintaining a constant temperature, is called a *thermostat*. The simplest switching thermostat works in this way: the T.S.E. is linked to a switch in the heater circuit such that the heater circuit is closed at ambient temperature, and as the body heats, reaches a pre-determined temperature at which the T.S.E. switches the heater 'off'. By loss to the ambient the temperature of the body falls, thus causing the T.S.E. to change and close the heater circuit, introducing more heat into the body and starting a second cycle of a repeated series of switchings. Provided that the temperature of the ambient does not fall so low that continuous operation of the heater cannot achieve switching temperature in the body, a degree of temperature control has been achieved, despite ambient fluctuations. It is also obvious that no control will be obtained if the ambient rises above the selected control temperature.

This system is, inherently, one of *control within limits* for the following reasons:

Thermostat differential—A mechanical system in which the T.S.E. directly opens and closes the heater circuit is converting the continuous movement of the T.S.E. into a make/break action at a defined point. Any appreciable heater current at a practical voltage cannot be interrupted or turned on by a slow-acting contact system; the T.S.E. must thus actuate a 'snap-action' switch. This is most frequently achieved by a bi-stable spring or by biasing with a permanent magnet; there is no direct mechanical method of achieving a q.m.b. switch from a continuously moving actuator without appreciable hysteresis, for the temperature at which the heater circuit is opened will be appreciably above that at which it is closed, and one is very lucky to get a commercial direct-switching thermostat in which this differential is even as small as $\frac{1}{2}^{\circ}\text{C}$. Further, since such a switch will need appreciable activation force, the T.S.E. must consist either of a substantial bimetal rod, or of a liquid-filled bellows, neither of which can (see Thermometers) be at all

TEMPERATURE CONTROL

sensitive. Where any attempt is made to reduce the mechanical differential, it usually results in a metastable condition of the contacts, producing continuous arcing, and the worst possible conditions for the operation of metal switch contacts.

Reduction of differential—The two principal methods of reducing the mechanical differential are by the use of mark-to-space systems (see below) or by the use of relays. As soon as the contacts operated by the T.S.E. are freed of the requirement to switch an appreciable heater current, the size and nature of the T.S.E. can be reduced to sensitive proportions, and if the relay current is sufficiently small the thermostat contacts may not need to open and close as quickly. However, it is still not possible to use the most sensitive contact thermostats directly with even the smallest solenoid-type of relay; they should be used with either bimetal or capsule type elements, or with very robust mercury-toluene regulators.

Ideal make-and-break relay—In the most refined systems of this type, the contacts of the T.S.E. operate through a single-stage valve relay, the anode of which controls a solenoid relay which in turn switches the heater current. This system has been used by the author with complete success, particularly for the control of aquarium tanks to limits of $1/50^{\circ}\text{C}$ (Figure 29.19). The contacts of the T.S.E. may be handling currents of a few microamps at a few volts, which can thus be interrupted by a minute movement of the element, however slow, without arcing, oxidation or similar attendant problems. The mercury-contact thermometer or sensitive mercury-toluene regulator is usually used in such systems.

Thermal hysteresis

It might appear from the above that, with the most refined system, immediately the required temperature had been reached and the heater

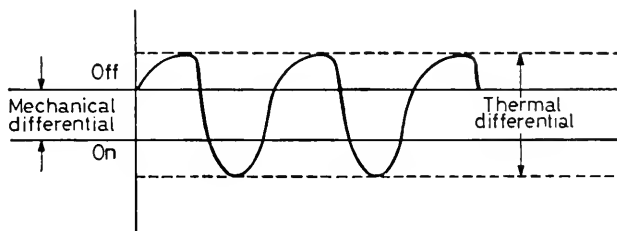


Figure 29.8 Typical form of thermograph record of crude switching thermostat, showing that the total temperature swing is greater than the mechanical differential of the switch

switched off the temperature will fall, and the heater immediately be actuated again, leading to a very rapid oscillation of the contacts (and hence of the relays). These conditions are in fact extremely difficult to achieve. A heater has a finite and often quite substantial thermal capacity; further, if it is to heat a body, the only way it can transfer heat to it is to be at an appreciable temperature above that body. The interruption of power to the heater causes, immediately, only a comparatively gradual decline in heater temperature, and hence of supply of heat to the body; thus the temperature

of the body continues to rise after the thermostat has acted, then reaches equilibrium, and *then* starts to fall towards a temperature at which the thermostat is closed. Correspondingly, the heater does not attain full working temperature immediately the supply is restored by the consequences of the body cooling. Thus the temperature of the body continues to fall, then reaches equilibrium, and only then passes into the 'rising' part of the cycle (Figure 29.8). This thermal cycle, then, has to be added on to such thermostat differential as exists, and with an unfortunate choice of components it can amount to several Centigrade degrees.

Heater design

The basic fact that a heater must be above the temperature of that which it heats must mean that there is always a temperature gradient from that part of the 'controlled' body nearest the heater, dropping in all directions towards the ambient. The higher the operating temperature of the heater, the greater will be the variation within the various parts of the controlled body. Conversely, the lower its temperature the less efficient will be its transfer of heat energy, into which process the thermal capacity of the heater material as well as its conductivity enter. Choice of the operating temperature of the heater may also involve economic considerations. Generally speaking, low-temperature heaters should be used with gases, while only higher temperatures (or very massive heaters) will be able to supply the heat quantity required by typical liquids and solids. Where fluids are force-circulated higher temperatures can be used, but if convective stirring predominates the heater should be at as low a temperature as is possible. Similarly, a form of phase shift is introduced by the accessory parts of the heater, and most especially by any covering over the element; these accessory parts will be heated, and will cool, at a different rate to the heater element itself; they are a mixed blessing. For gases, the secondary parts of the heater must be reduced to the absolute minimum, and a bare heater wire should make direct contact with the gas. For liquids, maximum conductivity from the heater may be achieved by, for example, supporting the heater wire by ceramic beads in an oil-filled metal tube or placing a spiral of wire so that it lines the inner walls of an oil-filled glass tube. It may be found that the 'smoothing' of the discontinuous heat supply by the thermal capacity of the heater enclosure does in fact assist in the close control of a large volume of water.

Heater-thermostat link

It is essential to realize that the *only* part of a body whose temperature is actually being 'controlled' is that in direct contact with the T.S.E. of the thermostat! The information on which the heater is being controlled is that presented *at* the thermostat; in general it is also here that we can expect the smallest fluctuation in temperature. We can in fact regard the control mechanism as an infinite gain amplifier to which the thermal link between the heater and the thermostat adds overall negative feedback. There is however an appreciable time constant in this link. It may be several seconds after the heater has come up to operating temperature that the temperature of the T.S.E. starts to rise, by which time an excessive amount of heat may

TEMPERATURE CONTROL

have been poured into the body in general; but the solution to this problem is not that of placing the T.S.E. very close to the heater, for if the body is large, the outlying parts of it may be hardly controlled at all. What is essential is to improve the thermal link by such devices as the forced circulation of the fluid medium (which will usually make up the body) as well as to ensure the maximum of contact between the T.S.E. and the material of the body. On the other hand, very vigorous stirring, especially of air, will introduce an appreciable heating effect in itself, especially if the stirring motor is inside the environmental enclosure.

Switching thermostat at different temperatures in its heater range

For a given fixed heater capacity, the range of available operating temperature will span from that of the ambient to that obtaining for the continuous operation of the heater; intermediate temperatures may vary from circumstances when the heater is operating for a minimal part of the time in one cycle of operations, to a maximal part of the time. From the foregoing

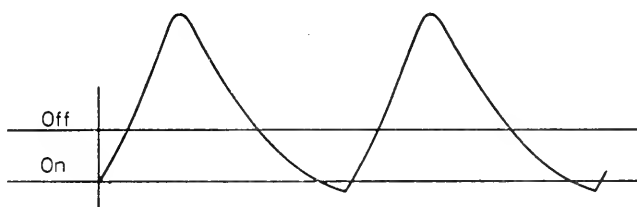


Figure 29.9 Thermograph record of a crude switching thermostat showing extreme temperature swing when working close to ambient temperatures

analysis it follows that the poorest control is obtained when operating at temperatures nearest to the ambient; cooling is slow and heat enters in short bursts of high level (see *Figure 29.9*), while at the highest levels of temperature the high cooling rate ensures that the effect of cutting the heat supply is rapidly accommodated, with the result that the frequency of a cycle of thermostat operation is much greater here. Though versatility of temperature range may be a requirement of a piece of equipment, it is obvious that the size of the heater should be most carefully chosen where one temperature is particularly required.

Split and multi-stage heaters

A great deal of the inherent temperature fluctuation in the switching thermostat system of control is thus due to the characteristics of the heater itself. Especially for higher temperature work, where comparatively large heaters are concerned, the heater may be split into one or more units with advantage. In the simplest system, especially where one particular temperature or small range is required, this takes the form of a continuously operating heater raising the body to a particular level, together with a smaller heater operated by the thermostat which 'tops up' the system. It will be noted that in large equipment this has the added advantage that the major part of the heater current does not have to be switched by the relays of the thermostat; on the other hand, the largest fixed heater which can be entertained is that bringing the body below the lowest temperature required

when the ambient is at its highest. In more exacting systems, where a wide temperature range is also required, the position of a T.S.E. may select an appropriate number of sections of 'fixed' heater, in addition to which either the same T.S.E. or a separate one may complete the topping up process.

Ambient compensators

For the largest installations, where economy of power becomes an essential, it is advisable to switch in extra sections of heater via a separate thermostat in the ambient, especially if this is subject to large fluctuations, and/or insulation is poor. Great care must be taken that this heat load can *never* produce temperatures which could overlap the working range of the internal thermostat, otherwise the most complex cycles of instability will be set up.

Mark-to-space systems

So far, control methods have aimed at minimizing the cyclical temperature fluctuations due to the method of heating—obtaining a fractional mean power from a discontinuously operating heater. The frequency of the operating cycle is a characteristic of the control machinery; in coarse control a cycle of several minutes is not unusual. Now approximately the same mean power emission could be obtained by any frequency of cycle, provided that the ratio of heater 'on' to 'off' remains constant (it is not exactly the same because the dissipated power also depends to some extent on the duration of actual heating and cooling of the heater element itself which will be greater in long frequency cycles); but the higher the frequency, the more even will the temperature control be. All the refinements outlined above do in fact tend to increase the frequency of the thermostat cycle, but a much more satisfactory method is by the use of a 'mark-to-space' system. The simplest of these is now dealt with.

Proportioning head—Consider (*Figure 29.10a*) the movement of the meniscus of a mercury-contact thermometer or toluene regulator; the 'on' and 'off' portions of the cycle are determined by a line representing the position of the needle contact. Now suppose the needle be made itself to oscillate vertically (*Figure 29.10b*) with a frequency appreciably greater than that of the mercury cycle. The sum of the 'on' periods per mercury cycle—and likewise of the 'off'—have not been changed, the mean power dissipation will not have altered but the distribution of power in time has been made very much more even. Because of this, of course, the amplitude of the mercury meniscus will be so materially reduced that the relative roles of meniscus and contact in determining the mark-to-space ratio will be reversed (*Figure 29.10c*). The much smaller movement of the mercury in itself indicates the improvement in control. Should the temperature of the body rise for any reason (e.g. rise of ambient) the 'on' period *per needle cycle* is at once reduced, resulting in rapid compensation. Perhaps the greatest advantage of the method is the ease with which it can be attached to existing mercury regulators; movement is usually obtained by mounting the needle on a bimetal strip wound with a small independent heater element whose circuit is broken by the strip's own movement, and remade as the strip cools again. It could be derived mechanically from a motor (but see next paragraph), and the application of the principle to electronic methods is of particular interest.

TEMPERATURE CONTROL

In order to obtain a new temperature setting the entire needle mechanism is of course moved to an appropriate new mean position.

Detailed consideration of mark-to-space systems—The size and form of the oscillation of the needle is critical. Compensation is due to the change in the mark-to-space—the ‘on-to-off’ ratio. Suppose the form of the contact oscillation is a sine wave, such as might be derived from a motor, and that the thermostat is set in its mid-range for the heater so that the on-off periods

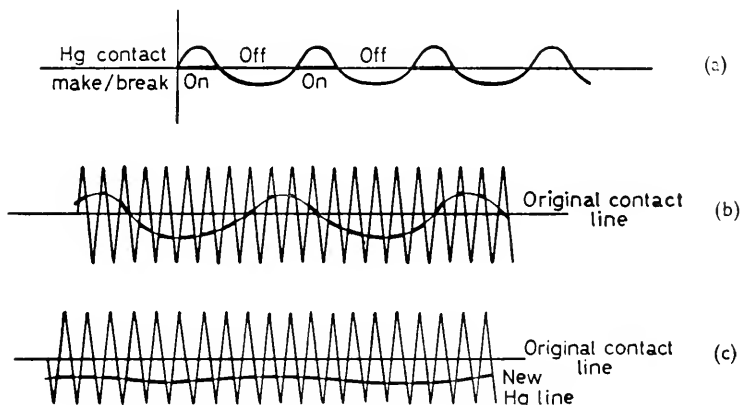


Figure 29.10 The effect of superimposing a proportioning head on a mercury-contact thermostat. (a) Unmodified thermostat; movement of the mercury meniscus (equivalent to temperature fluctuation) across the stationary contact position; (b) the contact is now oscillated by the proportioning head and its movement superimposed on the previous meniscal fluctuation, the intersection of meniscus and needle still determines the on/off periods, ‘on’ being indicated by the thickened parts of the curve; (c) condition of final control, the more even distribution of heat in time has achieved a much smaller temperature fluctuation

are about equal (*Figure 29.11a*). The intersecting ‘mercury line’ will have to move appreciably across this form of oscillation in order to change the ratio and produce compensation; the sensitivity will thus be low. If the apparatus is, on the other hand, working near to its upper or lower limits so that the oscillation curve is cut near its peaks (*Figure 29.11b*), the ratio will be rapidly changed by a small shift in the temperature line—so much so that over-compensation and instability may result. Uniform and adequate compensation over the entire available temperature range can only be obtained from a triangular type waveform (*Figure 29.11c*) which is not so readily obtainable. The wave front must be appreciably steep, but the more rapid within reason is the frequency of its oscillation the less need is there to approach the triangular form. However, one must still retain a sense of proportion about this, for presumably heater currents are being switched via relays, whose frequency of operation—and life—are limited.

Instability in mark-to-space control—Over-compensation may lead to instability—the regular form of large temperature cycles which may occur under these circumstances is called ‘hunting’. It can often be cured either by a reduction in the sensitivity of the T.S.E. or by increasing the amplitude of the hunting head; but in stubborn cases an analysis of the entire apparatus is often necessary. Theoretically (*Figure 29.12a*), the temperature of the

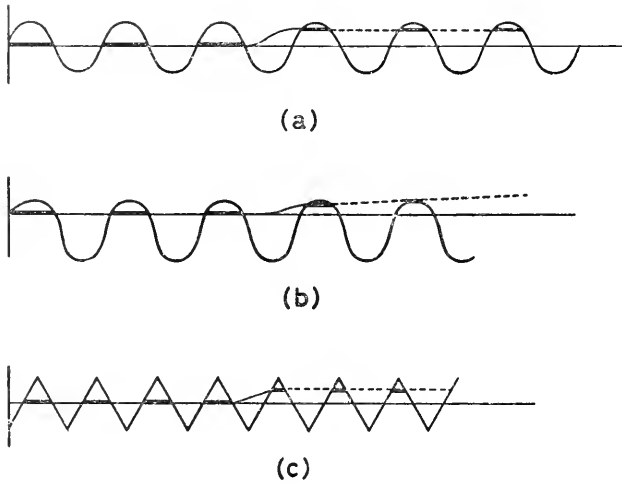


Figure 29.11 Forms of mark-to-space cycle. (a) Contact oscillating in a sine wave, a large temperature change gives rise to a small change in the ratio intercept, when operating centrally on the sine wave; (b) when operating near the apices of the sine wave, a small temperature change produces a large change in intercept ratio, (c) the change in the mark-to-space ratio is directly proportional to temperature change when a triangular wave is used

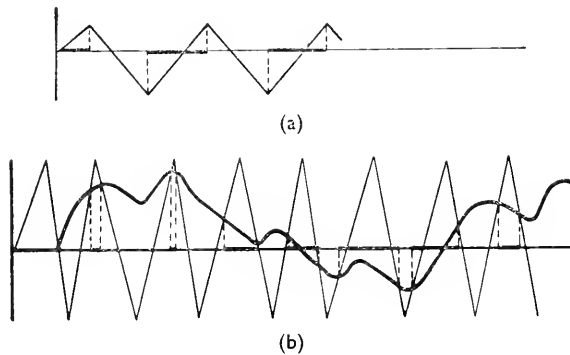


Figure 29.12 Effect of phase shift on temperature control by mark-to-space systems. (a) Ideal state in which temperature rises in the 'on' period and falls in the 'off'; (b) form of large-scale hunting due to 180 degree phase shift, i.e. a delay of half a mark-to-space cycle between the heater being energized and the corresponding temperature rise in the T.S.E.

TEMPERATURE CONTROL

body rises in the 'on' part of the cycle and falls in the 'off' part: this would certainly be true in the impossible condition of the T.S.E. being in contact with the heater; in fact there will always be a time-lag between the two. The thermal capacity of the heater might itself smooth out the intermittent nature of its power supplies, but especially in the most refined systems, where forced circulation assists the distribution of heat in a fluid, it is all too easy for the time lag to be of the same order as half a cycle of the hunting head. When this happens we have a form of positive feedback, and temperature fluctuations can become violent (*Figure 29.12b*). A further danger is present, particularly in air-bodies having a defined closed circulation (e.g. a plant breeding chamber); in one apparatus built by the author, where the air circulation and the mark-to-space cycle were both of the order of one minute, the cycles actually beat in and out of phase every two hours, giving a regular and large two-hourly temperature cycle. Whenever there are indications of longer term variations in a controlled system it is essential to discover by sensitive recording machines whether this is in fact cyclic in nature, and to suspect a positive feedback system if this is so.

Relays for mark-to-space systems—As the working of a mark-to-space system relies on the accurate division of a time interval it is clearly essential that there shall not be any variable time factors in the accessory parts of the mechanism. Thermostats of this kind are universally used with relays to switch the heater current, and while mechanical parts of relays will inevitably have a time constant this must not only be much shorter than the time cycle, it must be an absolutely constant delay, or it will modify the controlling interval. Mercury relays are particularly useful for this purpose. Relays of the hot-wire switch type are notoriously unreliable for three reasons: they are not only slow-acting, but have materially different periods for their opening and their closing movements; the variation in the energization to switching interval is very large; when operating with a frequency greater than 4 per minute the interval between energization and switching also depends on the previous switching experience.

In general, mark-to-space systems will put a greater strain on the relay system than a more simple type of control, and relays must be chosen carefully; the contact device of the thermostat must be guarded against arcing and literal burning, and there is everything to be said for the use of an electronic relay between the thermostat and the large mechanical relay (see *Practical Circuits*).

Non-linear temperature-sensitive elements

We have assumed in the foregoing that the T.S.E. has changed substantially linearly with temperature—that it has in fact been a typical thermometer element. Although they have not found use in very sensitive work, for a limited range of required temperatures there is an advantage in an element with a very great rate of change over only a small part of the temperature scale; thus a mechanical T.S.E., such as a metal rod, would have to be very long indeed to produce a big change in movement per degree, and even a bimetal spiral is still excessively high in thermal capacity to be at all sensitive. Likewise, one could use a Beckmann-type thermometer with inserted contacts, but the volume of liquid would be much too great for use with anything

except water as a medium. If one considers the critical vapour pressure of a sealed vessel of suitable liquid, there occurs a narrow range of temperature, when the liquid is vaporized, with a sudden large, and mechanically powerful change in volume. In practice, a bellows-type capsule of copper or bronze is filled with a suitable mixture of liquids, such as alcohol and ether, and sealed. Below a particular temperature the contents are at sufficient pressure to be entirely liquid; on heating a temperature is reached where rapid vaporization occurs, and under considerable pressure the capsule expands, often to many times its original length, and with appreciable force. Above this temperature again, only slight typical expansion takes place. The range covered by any one capsule usually amounts to a few centigrade degrees. It finds a most useful and efficient application in incubators where it operates heater switches directly: another advantage is that a new range is immediately obtained by substituting a different capsule, and the 'fine adjustment' of the system need only cover the few degrees of the individual capsule. On the other hand, though the capsule may appear small in size the energy it absorbs in operation is latent heat of vaporization, usually many times the specific heat of a liquid, so that in this respect it is an element of high 'thermal capacity'.

Approach characteristics

So far we have considered only the theoretical aspects of fixed-heater control when in operation at the controlled temperature. In many experimental devices the behaviour of the mechanism when it has had to depart from control is of considerable importance, e.g. the speed of recovery of temperature of an enclosed air space which has temporarily been cooled by opening a door on to the ambient; experiments in which material might be severely damaged by high temperature overshoot; equipment giving an alternating diurnal temperature cycle, where the temperature must rise and fall between two controlled levels at pre-determined times. For fixed heaters and simple thermostats the heating rate during approach to the controlled temperature is clearly defined by the size of the heater (*Figure 29.13a*), unless one is operating very close to the maximum temperature obtainable by the heater. If a very rapid return after intermittent cooling is an essential requirement of the equipment it may be necessary to have a heater much larger than would be used for maintaining the desired temperature; there would consequently be poor regulation. These circumstances call for a split heater, of which the larger portion, a booster, is switched *out* when the thermostat approaches the control temperature. Even so, the thermal capacity of the booster heater and its supports may lead to overshoot unless it is switched off sufficiently early (*Figure 29.13b*).

The behaviour of control gear of an enclosed air space while it is temporarily open to a cool ambient must be closely watched. Regardless of the thermostatic arrangement, this will usually lead to all heaters being fully energized; if opening the box also interferes with the circulation system there may be violent local heating.

The form of approach in single heater equipment with overshoot is often followed by a typical train of damped oscillations (*Figure 29.13c*) until control is obtained—this is not a danger sign of instability. As a rule the

TEMPERATURE CONTROL

mark-to-space type of control approaches control from lower temperatures smoothly, and without overshoot (*Figure 29.13d*), but, in general, the more sensitive the equipment the slower will be its approach to the control temperature; if the sensitivity is increased to the brink of instability a large disturbance (which is the equivalent to meeting approach conditions) will lead to instability, and it may be advisable to have a variable sensitivity control which is brought into action during such operations as opening doors. Where equipment is needed to alternate in temperature on, for example, a diurnal

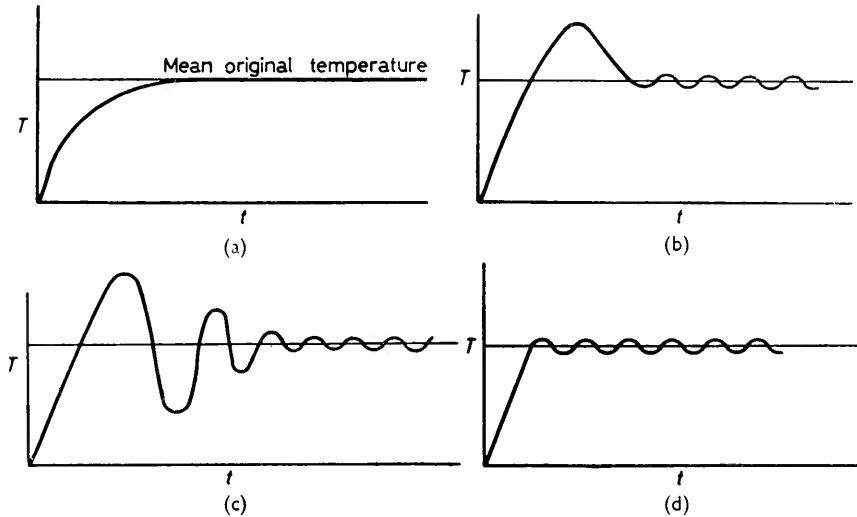


Figure 29.13 Approach characteristics of controlled bodies heated from a lower temperature. (a) Asymptotic approach of a body to the maximum temperature produced by a fixed wattage heater; (b) overshoot caused when a booster heater of large thermal capacity is not switched off till control temperature is reached; (c) 'damped oscillation' produced by a thermostatically controlled single heater of large thermal capacity; (d) approach of a small thermal capacity heater, controlled by mark-to-space—heating is at maximum rate until control temperature is reached, followed immediately by optimal control

cycle, we must also consider approach to a lower temperature from a higher one at the switch-over point. This is, in general, regulated by the insulation of the body from the ambient, and as it is not impossible to construct an environmental chamber which will lose only a few degrees per hour with its heater switched off, it follows that for such equipment one should have only the minimum of insulation compatible with the necessary cooling rate, bearing in mind that the heater is going to have to be appropriately bigger, and that the machine will be less economic to run. Provided the equipment is always required to run above ambient temperature it is not advisable to introduce active cooling for this (see below).

Continuously variable heater system

In all the foregoing the effect of the control apparatus has been to distribute the action of a fixed-power heater in order to obtain the required mean power necessary for a temperature lower than that produced by continuous

operation of that heater. An entirely new concept of thermostatic control arises if the heater could be of continuously variable power output—a condition which one attempts to produce in mark-to-space systems with the most efficient stirring of fluids, and a high thermal capacity of the heater itself acting as a smoothing agent. A second, more crude, approach is obtained in the largest installations where multiple-contact thermostats control a succession of heaters, and a high degree of control is impossible. Continuously variable control of heater power output is however achieved in the method of Beament and Machin¹. The heater resistance is made the anode load of a power valve, and its heat output is thus controlled by the grid voltage of the valve. The information to be fed to this device can no longer be of the 'on-or-off' variety, for the valve requires quantitative information of the amount of heat required to control the body's temperature; the thermostatic element must therefore establish a reference voltage and vary this in a suitable sense with temperature fluctuation in order to obtain compensation control. A resistance thermometer is clearly the most suitable T.S.E. for this purpose and a thermistor has proved excellent in practice. It is placed in a potential divider such that a change in its resistance controls the power output of the valve in the appropriate sense. The sensitivity of the system is increased, and is made variable by positive feedback, this loop being within the negative feedback loop of the T.S.E. to the heater itself. The device gives a large swing in a continuous power output to compensate for small temperature changes. This is probably the most accurate fine control for air temperatures yet devised, but although it has no inherent 'differential' in the sense of the discontinuous heater it is obvious that it can only compensate because the temperature of the body *does* change. For best results, everything which has been said about good insulation, circulation, heater size, etc., must be adhered to.

Continuously variable temperature

Theoretically speaking, there is no greater problem in making a body follow a temperature programme than maintain a desired temperature. All that is needed is to ensure that the maximum *rate* of temperature increase and decrease can be obtained, and to programme the temperature setting unit of the thermostat. Practically speaking, such a requirement will need complex controlling equipment and the most careful design; it should be avoided if at all possible.

Control below ambient

There is no basic difference between using a thermostat to control a cooler rather than a heater—a refrigerator does this crudely as an everyday affair. The removal of heat energy is a more difficult engineering problem, and from a thermostatic point of view one which is slow and cannot be switched rapidly on or off, or made quickly variable. Practical methods will be outlined below; no new principles are involved.

Control both above and below ambient

This can be approached in three ways: the entire ambient can be cooled below the lowest temperature needed, and then conventional control of

PRACTICAL METHODS FOR TEMPERATURE CONTROL

positive heating used—this is obviously very expensive unless the installations are small. The body may be provided with a continuously operating heater, or cooler, with control of the other factor—the most usual way of doing things, though again this is extravagant of energy; where economy is essential it is better to have continuous heating with control of cooling, though the reverse process leads to the closer control. The dual control of heat and of coolth is fraught with every possible snag. This is especially true if total heater and cooler switching is used; either there must be a point at which one switches from heater to cooler or there will be a range over which both heater and cooler are operating simultaneously—and no pack of hounds will ever hunt as efficiently. The author has overcome this problem in a machine controlling the temperature of circulating fluids by using a special mark-to-space system described in the next section; it saves 50 per cent energy at intermediate temperatures, and can move from the high temperature of continuous heating to the low point of continuous cooling. The device is mechanically complex but is particularly useful for such problems as plant-breeding in the laboratory.

External load changes

Of the special problems arising in constant temperature enclosures, the commonest is the influx of energy through light. The sort of light intensity needed to breed plants may easily introduce energy causing a rise in temperature of several degrees. Some workers have chosen the simple method of including a fixed heater, which is switched in as the lights are cut, and of a value to keep the heat load constant: as a rule, large-scale cooling is needed continuously to off-set this. The practice seems excessive, and it is far better to place all such lights outside the enclosure and employ every possible means of keeping down the entrant energy, which will be especially heavy in the infra-red. A well designed thermostat should then meet the situation adequately.

PRACTICAL METHODS FOR TEMPERATURE CONTROL

It is assumed that the reader is especially interested in the control of temperature of air or water containing biological material, and that ovens, incubators and autoclaves of conventional design will usually be purchased ready made.

Aquaria and water-baths

Although economically lagging may be worth while, and is advised where such enclosures are run appreciably above or below ambient or where the tank is of metal, it is perfectly possible to obtain temperature control to $1/50^{\circ}\text{C}$ in an unlagged glass tank; however, for normal purposes, control to $\pm\frac{1}{4}^{\circ}\text{C}$ will be adequate. The heater should consist of a spiral of nichrome or tungsten wire in a glass tube filled with liquid paraffin, supported off the floor of the tank. A simple stirrer should be included, but a stream of compressed air through the water is often sufficient to prevent stratification—the density change with temperature is such that convection in water will not lead to large temperature variations as compared with air. An indication of the required size of heater may be obtained by putting in a small immersion

MEASUREMENT AND CONTROL OF TEMPERATURE

electric heater on a basis of 10 W per litre and leaving the tank (stirred) to obtain its highest equilibrium temperature. A rough calculation based on this information (W lost per hour per temperature unit above ambient) will enable one to obtain the size of heater for the temperature range required. Up to 50 W might be switched direct on a bimetal thermostat (*Figures 29.14*

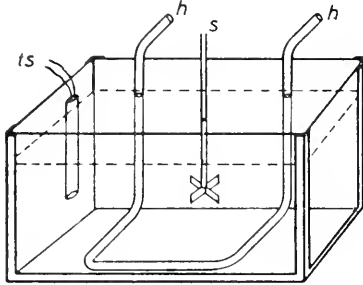


Figure 29.14 Layout of tube-enclosed heater (hh), stirrer (s) and thermostat element (ts) in an aquarium tank

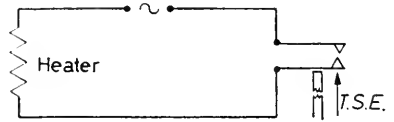


Figure 29.15 Simplest thermostat circuit in which the T.S.E. switches the heater directly

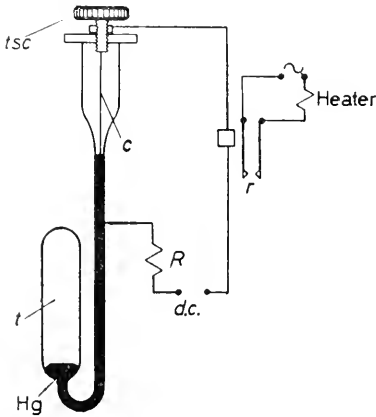


Figure 29.16 Schematic layout for control by robust mercury-toluene regulator (t, Hg). Temperature setting by the insulated screw contact (tsc) and relay (r) current adjusted by resistance R

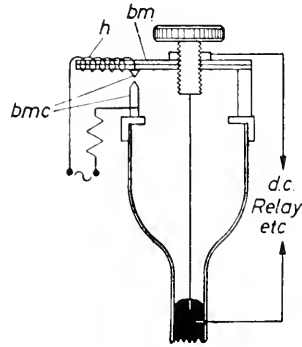


Figure 29.17 Schematic form of proportioning head on mercury-contact regulator. The adjusting screw is mounted on a bimetal strip (bm) heated by the winding (h) energized through the contacts (bmc) which are opened and closed by the strip's own movements

and 29.15). Above this load, and for finer control, a relay should be used, or, better, the relay should be controlled through a robust mercury-toluene regulator (*Figure 29.16*). The addition of a proportioning head (*Figure 29.17*) will obtain control to better than $1/10^{\circ}\text{C}$.

Temperatures below ambient in small installations are best obtained by running a cooler coil from a refrigerator exchanger; glycerin/water is advised for the circulating fluid as it is non-corrosive. For multiple tank systems do

PRACTICAL METHODS FOR TEMPERATURE CONTROL

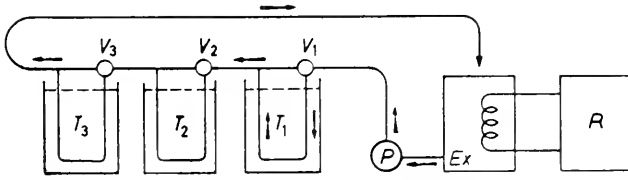


Figure 29.18 Essential layout for control by circulating cold liquid. The refrigerator (R) controls the temperature of the exchanger tank (Ex) from which circulating fluid is pumped (P) to the controlled tanks (T_1 - T_3) in series. At each tank a solenoid-operated valve (V_1 - V_3) can switch the liquid either through the tank or through the by-pass. Each solenoid is operated via a relay and the thermostat in the appropriate tank; tank T_1 will run coolest

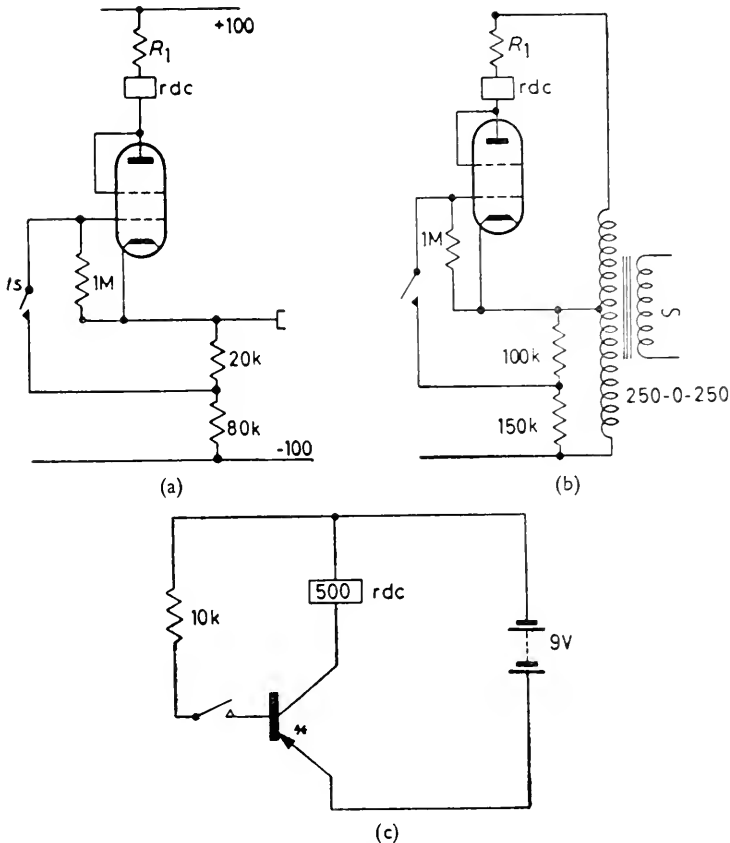


Figure 29.19 Electronic relays for use with very sensitive contact thermostats; in all cases the sense of the relay must be chosen (normally open or closed) in relation to the control of heat or coolth. (a) A large tetrode operated from 100-0-100 V d.c. mains; the relay (rdc) and its series resistance R_1 must be selected according to the valve, but for an 807 valve, a 32 mA, 50 V relay is suitable. (b) Similar circuit using unrectified power supply; the power available at the relay will be only about $\frac{1}{3}$ that of a corresponding d.c. plate voltage and the relay must be selected for anti-chatter. (c) Transistor operated relay designed for a P.O. relay. Almost any small junction transistor is suitable

not attempt to run cooler spirals in parallel—run them in series, with the coolest tank nearest the outlet of the exchanger reservoir, for the problem of maintaining many flows in parallel from one pump is extremely difficult. For large installations it will prove much more economic in terms both of size of refrigerator unit and also of refrigerator running costs if one has constantly maintained heaters and controls on the refrigerator circulating fluid (*Figure 29.18*) by a solenoid-operated valve. It will be noted that the thermostat now is required to introduce *cooling* on the 'make' part of the cycle.

Very fine limits of control are obtained by using a mercury-toluene regulator, or mercury-contact thermostat, operating through a single stage electronic relay which in turn energizes a mercury relay switch—operating the eventual heater or cooler as the case may be (*Figure 29.19*). Almost any power triode, tetrode or pentode can be used, but there is a great deal to be said for using a very robust valve operating well below its rated anode voltage, in which case its life will be almost indefinite.

There is the further advantage that in laboratories where 100–0–100 V d.c. is available such a unit can be run without a special power pack; alternatively, germanium rectifiers could be used to obtain suitable plate currents.

Air enclosures

The most frequent biological requirement is the plant-animal breeding chamber. More variations in design and principle of these rooms and boxes have been tried—and indeed are in existence—than in any similar piece of laboratory equipment setting out to achieve one object, the provision of constant air temperature with additional lighting equipment.

Insulation is of great importance, and at least two in. of cork, slag wool, expanded plastic foam or similar material is needed, which must be securely sealed into water-tight compartments, for moisture degrades insulation remarkably. An indication of the efficiency of insulation can be gained from the tests which should be carried out to determine heater size; 5–10 W per ft.³ of air should be sufficient to maintain the enclosure at 20–30°C above ambient. An ordinary tungsten filament lamp (which produces 95 per cent of its rated energy consumption as heat) will suffice as a test heater. The eventual heater (*Figure 29.20*) should be wound of nichrome wire, so that air will circulate *through* the wire. For simple systems asbestos 'blanket' heaters are ideal. The heater should be so shrouded that the circulating air is positively passed through it. A low-pressure high-volume fan is needed to provide good air circulation—remembering that if the motor is inside the enclosure it also will dissipate its rated consumption in heat, either directly or via the energy of stirring. For temperatures below ambient where refrigeration is used it is again essential to blow the circulated air through a large surface area radiator; unless a special mark-to-space method is adopted (see below), it is not recommended to control on the cold side. Usually one would place cooler and heater in the same shroud so that the air is cooled, and then heated under thermostatic control; care should be taken that condensation on the cooler does not drip on to the heater.

It is appreciated that the temperature of air in an incubator can be held to reasonable limits (say $\pm \frac{1}{4}^{\circ}\text{C}$) using a relatively crude thermostat; incubators normally require many hours to come into equilibrium; they are slow to

PRACTICAL METHODS FOR TEMPERATURE CONTROL

recover from such disturbances as opening doors, and they usually rely on having a water-jacket or similar surround of high thermal capacity to buffer changes and to slow cooling by buffering the insulation. In typical experimental enclosures it is usually necessary to have an observation site (for which double glass or plastic walls, with at least an inch of air between, are necessary), a window for admission of light, and a door which may have to be opened frequently; most people would hesitate to build a water-jacketed box with these attributes, while of course high thermal capacity walls are

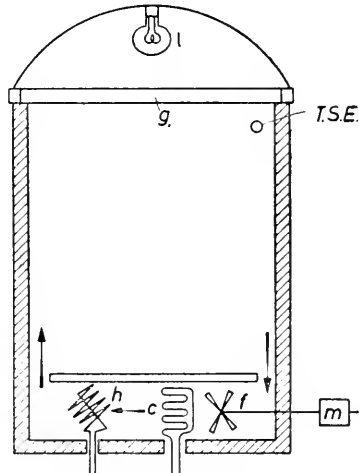


Figure 29.20 Suggested sectional layout of a plant-growing cabinet. The heater (h) and cooler coil (c) are shrouded below a false floor through which the air is blown by a fan (f), the drive motor (m) being outside the box. The lighting equipment (l) is in an aspirated compartment separated by two sheets of glass (g)

directly contrary to a requirement of frequent shifts in actual controlled temperature. Generally speaking, only a small mercury-toluene thermostat or, better, contact thermometer will meet these circumstances, together with an electronic relay operating an electromagnetic relay with mercury contacts. It is true that some not unsuccessful units operate on the principle of placing heaters and/or coolers between a false wall and the main insulation jacket, so that the air is 'enclosed in a box at the right temperature'. Such an extension of the incubator principle is likewise useful for long usage at one fixed temperature, but slow in response to change; it is also more difficult to construct. If such a system is used for low temperatures, heavy condensation will occur on the walls and may be a great nuisance.

Large-scale mark-to-space systems

The author has used the system outlined in *Figure 29.21* for a plant-breeding cabinet. It uses circulated air, switched by a mark-to-space thermostat either through a heater or through a cooler, using a solenoid-operated air valve. Its great advantage is of rapid accommodation of

disturbances, together with the economic consideration that for low temperatures cooler alone is principally used, while for high temperatures there is little energy drain on the cooler; further economy can be effected by providing a by-pass so that except on such occasions as the opening of the cabinet doors the major part of the air is re-circulated without reconditioning; humidity control can be independently incorporated on the same principle. The machine can be operated by mechanical or electronic switch-gear, and is particularly suitable for diurnal or continuously variable cyclical operations.

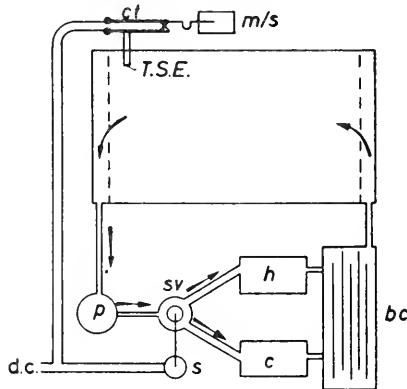


Figure 29.21 Layout for fine control of circulated air (large-scale equipment) by mark-to-space control, suitable for temperature and/or humidity control above and below ambient for fixed or programmed characteristics. Air drawn from the cabinet by the pump (p) passes through the solenoid-controlled valve (sv) either to a heater (h) or cooler (c) and into a baffle-chamber (bc) for mixing before re-entering the enclosure. A humidifier/drier unit would run in parallel with these. The control thermostat (ct) and mark-to-space injector (m/s) can be sited appropriately at the critical part of the enclosure

Its accommodation is such that it can readily be made to swamp the load changes of discontinuous illumination.

Fine control by variable heater

The circuit for the method of Beament and Machin, using thermistors, is given in Figure 29.22. The heater—2 k Ω of nichrome wire wound on an open frame of mica sheet—is supported in a cylinder of insulating material, and above it a small fan attached to a pot magnet is mounted on watch-maker bearings. A long bar magnet rotated outside the enclosure spins the fan without introducing energy from the heating of the motor or requiring a shaft through the wall of the box. The high resistance thermistor forms a potential divider to control the grid of the first stage valve, and temperature settings are obtained by adjusting the voltage of the upper end of the thermistor. An additional range is made available by switching out part of the resistance network. Two miniature neons provide a voltage drop to the directly coupled grid of the output stage, of which the heater is the anode load. A potentiometer in parallel with this load provides a proportional voltage which is put on to the screen of the first stage, thus giving variable positive feedback and allowing the sensitivity of the amplifier to be adjusted

PRACTICAL METHODS FOR TEMPERATURE CONTROL

to the required amount. Stability is of course provided by the overall negative feedback from heater to thermistor, via the air in the enclosure. The apparatus works quite satisfactorily with power supplies merely rectified by germanium rectifiers and smoothed by $32\ \mu\text{F}$. If, however, the sensitivity control is turned up so that the output stage is either turned 'on' or 'off' by movement through a small temperature excursion, then the heater can be replaced by a

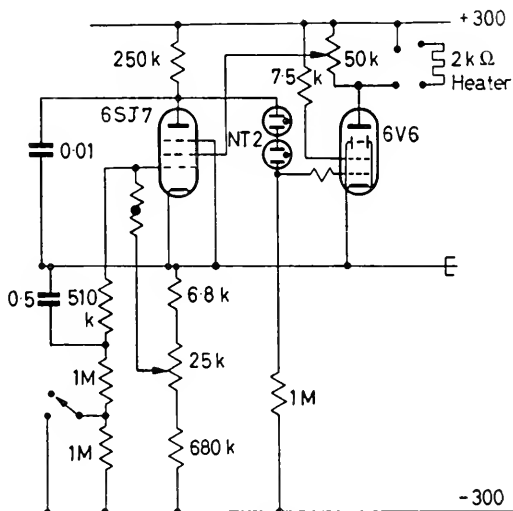


Figure 29.22 Circuit for ultra-fine control of air temperature in small enclosures; the thermistor is of type $100\ \text{k}\Omega$ at 20°C . Two control ranges are available from the switch which could also incorporate a background heater on the higher range

powerful relay which will control large heat loads. With this method, control of air temperature to $1/50^\circ\text{C}$ in a litre of air in an uninsulated plastic box has been obtained.

Hot stages

The temperature control of specimens usually immersed in liquids on the microscope stage has received much attention. The greatest problem associated is that of allowing the illumination of the specimen. Stages using circulating water in a transparent cell, and electrical heating of a perforated metal block, are in regular use. The important consideration here is that of putting the specimen in contact with a large reservoir of heat so that it can readily obtain the heat it itself loses by conduction or evaporation. Wherever possible the specimen should be supported on the minimum of insulating material, e.g. on a thin coverglass, and the surface which heats it must be polished to ensure the maximum of contact.

q.m.b. (page 394) means quick-make and -break, i.e. 'snap' or toggle action.—ED.

MEASUREMENT AND CONTROL OF TEMPERATURE

REFERENCE

BEAMENT, J. W. L. and MACHIN, K. E. *J. Sci. Inst.* In press (1958)

BIBLIOGRAPHY

The books and articles hereunder are almost entirely devoted to industrial measurement and control problems.

Temperature: its measurement and control in science and industry (1941) *Amer. Inst. Phys.* Reinhold

GRIFFITHS, E. G. *Methods of measuring temperature* Griffin. 1947

WEBER, R, L, *Temperature measurement and control* Blakiston. 1941

MEASUREMENT AND CONTROL OF HUMIDITY

J. W. L. BEAMENT

INTRODUCTION

The biologist is frequently interested in obtaining a relatively crude measurement of air-moisture content; he may find occasional need to employ crude control of the moisture in a plant-growing chamber or experimental enclosure, which usually amounts to obtaining a 'high' humidity or literal saturation of the air. However, the accurate assessment of air moisture content is an exceptionally difficult—or laborious—business while its control to anything like fine limits is therefore correspondingly difficult and usually has awkward inter-actions with fine temperature control, which will naturally be a prerequisite of such an experimental device.

UNITS

Relative humidity (R.H.) is the percentage water in air of that which the air would contain if saturated *at that temperature*. Saturation deficiency (s.d.) is the difference between the saturation vapour pressure of water at a temperature and that which actually exists; it is usually expressed in mm Hg. This latter is the most useful biological expression of humidity, for it indicates both the 'evaporative power' and the 'carrying capacity' of air, and it is further independent of temperature, whereas a percentage R.H. is only useful information when the temperature is also quoted.

MEASUREMENT

Fundamental consideration

A hygrometer may operate by indicating the rate of evaporation into the air; in this case it is continuously increasing the humidity of an enclosed space itself. Since evaporation automatically involves latent heat loss and cooling, the hygrometer must properly be provided with a compensating heater which exactly brings it back to ambient temperature—this would constitute an excellent problem in feedback engineering. In addition, the rate of evaporation will be entirely dependent on the rate of air-flow over the element.

A second method uses the dew-point principle: to cool a surface down to the temperature at which the moisture content saturates the air and therefore causes the surface to mist; the temperature difference and the true ambient temperature then become transferable into humidity units by the use of suitable tables. For conditions in which accurate temperature control is

being attempted this system is useless. For large-scale meteorology, machines with built-in controlled cooling, and photocells which detect the moisture film have been used with appreciable accuracy.

Thirdly, the 'wet bulb' temperature can be used—the temperature to which water is cooled by evaporation into the air. This is altogether both inaccurate and difficult in use.

Finally, there are methods in which materials are selected whose particular measurable property changes with water absorption. This may be the contraction and expansion of paper, hair, catgut, etc., the degree of ionization of a salt, usually as the conductivity of the electrolyte, the resistance of various porous materials in different states of hydration, etc. These are generally speaking the best methods of measuring humidity. They all suffer from different disadvantages: in particular that they are slow to come into equilibrium, and even more important that they cannot therefore follow rapid changes in humidity, that they have very pronounced hysteresis, and that their calibration changes with age and conditioning; that they are excessively susceptible to contamination of all kinds—particularly to minute traces of grease. They all to a certain extent will lose, or take up, water in the process of coming into equilibrium, though if the element is made small enough this can be reduced to small proportions, and once equilibrium has been obtained they will at least indicate (a) that a particular humidity is remaining constant and (b) that it has departed from that state to a certain extent. Many small portable measuring instruments such as paper and hair hygrometers have been marketed, in which the movement of the element is directly shown by a graduated scale and pointer, and have limited application over a particular range. This brief account will of course be limited to electrical applications to such instruments.

CALIBRATION

The calibration of hygrometers presents an extremely difficult problem and since the only arbitrary standards which can be reasonably achieved are 'dry' air and 'saturated' air the fundamental practice is to produce reservoirs of these (e.g. by drying through a liquid air trap and by bubbling through a sinter-glass filter in distilled water) and to mix in suitable proportions to give the required humidity *in a vessel which does not adsorb* water vapour—such as polystyrene, p.t.f.e., or silicone-coated glass.

As a reasonable laboratory alternative, tables exist of the relative humidities of air in equilibrium with saturated solutions of a large number of salts at selected temperatures. So long as care is taken to ensure accurate temperature control and purity of solutions this method will probably provide calibration humidities as accurately as any biologist may need.

PRACTICAL DETAILS OF ELECTRICAL METHODS

Thermocouples

The two junctions of a thermocouple circuit can conveniently be used as the wet and dry bulbs of a hygrometer; since they can be made with minute thermal capacity, the wet bulb evaporation can be very small, and

PRACTICAL DETAILS OF ELECTRICAL METHODS

will need little ventilation. The smallest wick is needed, together with a tiny reservoir. The device can either be used calibrated against the direct reading of the galvanometer or, alternatively, the galvanometer can be zeroed by applying a current to a small heater wound within the wick, and the machine calibrated against the heater current (*Figure 30.1*). The sensitivity can be greatly increased by using a thermopile of many junctions

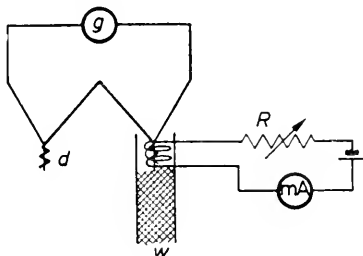


Figure 30.1 Compensated wet bulb system for thermocouple hygrometer. The wet bulb (*w*) has a small heater coil wound within its wick, and the current in it is adjusted until the galvanometer (*g*) is zeroed, readings are then made off the milliammeter

(see Chapter 29) for the two sets of junctions need not be greatly separated in distance.

Resistance thermometers

If the wet and dry bulbs are two resistance thermometers which form two arms of a bridge current, the out-of-balance will indicate a function of humidity; the law will be complex, and temperature dependent, but if a second similar bridge is so arranged that it contains a 'dry bulb' with a fixed

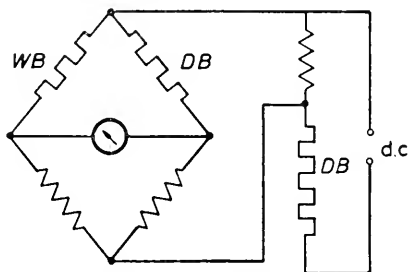


Figure 30.2 Schematic layout for controlling the voltage on a wet (*WB*) and dry (*DB*) resistance thermometer bridge, using a second dry bulb element

resistance in place of the wet bulb element, a suitable comparison of the outputs of the two bridges can be made to give a reasonable measure of humidity, independent of temperature, for the second dry bulb compensates out the first one. This can either be done by a cross-coil meter movement, which presents the resultant of the two bridge generated currents, or by using the output of the compensator bridge as the feed of the primary bridge, in which case the out-of-balance will give a direct measure of saturation deficiency (*Figure 30.2*).

Thermistors—sealed

In the same way that normal resistance thermometers can be used as wet and dry bulb instruments, so also thermistors of the glass encapsulated type can be used in a bridge. Because of the minute power dissipation which can be allowed with them (see previous chapter) the author has found it necessary to use an amplifier after the output of the control thermistor in order to generate a suitable voltage for the main bridge network. Certain ceramic materials containing metal oxides, which are thus essentially similar to thermistors, are also manufactured which can be used naked, and whose resistance is a function of the degree of adsorption of water—in itself a function of the surrounding humidity. The amount of water exchanged by these elements is very small, and if they are used to control the grid of a valve-amplifier the output of this will operate a meter, recording device, or be available for control purposes. They are, however, to be used with the utmost precaution as far as contaminants are concerned, and should be re-calibrated fairly frequently.

Electrolytic types

These depend on measuring the conductivity of a salt, typically lithium chloride, calcium chloride, etc., according to the range of humidity to be measured. In order to present the maximum surface for equilibration with the atmosphere, to retain the whole of the ionic material, and to obtain the highest possible overall resistance for a given amount of salt, the element is usually an appreciable length of 'thread' which has been dipped in solution and then dried by a good agent such as phosphorus pentoxide. An a.c. bridge technique must of course be used for measuring the resistance of the element to avoid polarization and electrolysis. The response time is very rapid but the device is complexly temperature sensitive, and must be re-calibrated for new temperatures. It should be possible to adapt this method to micro-measurement by the use of a very fine small elements, but they would be correspondingly delicate.

CONTROL OF HUMIDITY

In the same way that the control of temperature (Chapter 29) may involve the addition of suitable quantities of heat, or their extraction, or variously of both, so also humidity is controlled by adding or extracting moisture. On the other hand, whereas air may be readily heated, humidification is a more difficult proposition. It is meaningless to control humidity (except to 100 or 0 R.H.) without good temperature control also, and almost all humidification methods will produce thermal disturbances; conversely, temperature differences will produce humidity disturbances, and of course cool surfaces, especially refrigerated ones, will condense moisture and so lower humidities. However, unlike heat, it is perfectly practical to 'put' the right humidity in a box—and there it will stay provided the box does not adsorb water or change in temperature. Unhappily, the most likely source of interference will be through living material, especially plants, in the enclosure, so that compensation mechanisms have to deal with living—and therefore also will possibly compensating—mechanisms.

BIBLIOGRAPHY

Humidifiers

Air is not efficiently humidified by being blown *over* water; it can be most effectively humidified by bubbling through an immersed sinter-glass funnel or by being blown through meshes across which water is flowing. It is essential to use distilled water. Steam injection involves great quantities of heat release from condensation; the essential to accept is that it is useless merely to give air the opportunity to take up water—it must *be* humidified.

Driers

Apart from refrigeration condensation drying, which has every kind of disadvantage, about the only available drying process is chemical adsorption, and silica gel is the only readily recoverable chemical drier. It is quite sufficient for biological work and gives an excellent indication of its active life when impregnated with cobalt salts.

Control

Humidistats using every variety of element and principle are in commercial use. Typically, a mechanical element may be used such that a change in its length operates a pair of contacts, thus energizing a moisture, or drying, source through a relay. The nature of the time constant of these systems is very variable, but is typically long; they have great hunting instincts. Wherever possible mark-to-space systems, utilizing the humidity equivalent of the temperature regulating circuits shown in Chapter 29, should be used.

Capacity methods

Commercial moisture-measuring devices exist, basing their measurements on the relation between capacity and moisture content. They have excellent applications to the routine analysis of uniform samples, such as of grain, fabric, etc., but of course, a change in every constituent of air, and also both its temperature and pressure, will also alter its dielectric constant. It is therefore not a reliable method of measuring the humidity of an air sample which may vary in respect of temperature, barometric pressure, carbon dioxide, or other gas content.

Conclusion

This account of humidity measurement and control is necessarily brief because of the paucity of research and development in it in biological work, coupled with the technical difficulties and expense of such projects. Biologists are sincerely advised to regard all measurements of humidity with suspicion, to enter the field of humidity control as though it were at least the eighth circle of the Inferno, and to realize that however near to a living surface they may attempt to measure or control, the living surface will always have the last word.

BIBLIOGRAPHY

For references to the authoritative work of Professor H. S. Gregory in this field, see:

GREGORY H. S. The measurement of humidity *Instrument Practice* (1947)
— and ROURKE, E. *Trans. Soc. Inst. Tech.* 4 (1952)

MEASUREMENT AND CONTROL OF HUMIDITY

For industrial methods see:

MARSH, M. C. *Controlled Humidity in Industry* Griffin, 1935

For tables of constants see:

Hygrometric Tables: Meteorological Office report No. 265. H.M.S.O.

WILSON, R. E. Humidity over sulphuric acid solutions *J. Ind. Eng. Chem.* 13 (1921)

O'BRIEN, F. E. M. Humidity over saturated solutions of salts *B. Elect. Allied Ind. Res. Ass. Pub. L/T 166*

ASSAY OF RADIOACTIVITY

R. D. KEYNES

Reference books

It is impossible in a brief chapter to present more than an outline of the methods used for determination of radioactivity. There are a number of standard textbooks which cover the whole field of tracer technique, both in general and as applied to biological problems, among which those by Kamen¹, Sacks^{2,3}, Siri⁴, Whitehouse and Putnam⁵, Taylor⁶ and Sharpe⁷ may be recommended to the reader seeking a more complete account than can be given here. The Atomic Energy Research Establishment's catalogue of radioactive materials and stable isotopes (Catalogue No. 4 was issued in January 1957) also contains much useful information.

Isotope characteristics

The characteristics of the radiation emitted when radioactive atoms decay differ widely from one isotope to another, and play an important part in determining the most appropriate technique to be used in assaying any given isotope. Three types of radiation may be encountered: α -particles, which are helium nuclei with a double positive charge moving relatively slowly (order of 10^9 cm/sec), β -particles, which are electrons (or positrons) with a single negative or positive charge moving about ten times faster, and γ -rays, which are uncharged quanta of electromagnetic radiation moving with the speed of light. The various isotopes differ not only in the type of radiation that they emit (many of them produce β - and γ -radiation at the same time, though some are pure β - or γ -emitters) but also in the energy of their radiations, i.e. in the amount of energy dissipated when the process of radioactive disintegration of their nuclei takes place. *Table 1* lists the characteristics of some of the radioactive isotopes most likely to be needed in biological tracer experiments, and it will be seen that β -particle energies range from the very weak 0.018 MeV radiation of tritium (^3H) to the powerful 3.6 MeV β -particles given by ^{42}K , while γ -ray energies are distributed over a somewhat narrower range. The rate at which radioactive isotopes decay, in other words their half-lives (the time taken for half the atoms initially present to disintegrate), also varies widely. Some are so unstable that their half-life is a small fraction of a second, while at the other end of the scale ^{40}K , the naturally occurring radioactive isotope of potassium, has a half-life estimated as 4.5×10^8 years.

INTERACTIONS OF RADIATIONS WITH MATTER

In penetrating through matter the charged particles (α and β) interact electrostatically with the orbital electrons, and less often the nuclei, of those

ASSAY OF RADIOACTIVITY

TABLE 1

Characteristics of the Radioactive Isotopes of Greatest Interest in Biological Research

Isotope	Half-life	Energy of radiation in MeV	
³ H	12.3 y	0.018 β^-	no γ
¹¹ C	20.4 m	0.98 β^+	no γ
¹⁴ C	5,600 y	0.155 β^-	no γ
²² Na	2.6 y	0.54 β^+	1.28 γ
²⁴ Na	15.0 h	1.39 β^-	1.37 and 2.76 γ
²⁸ Mg	21.4 h	0.46 β^-	0.03 to 1.35 γ
³² P	14.3 d	1.71 β^-	no γ
³⁵ S	87.1 d	0.167 β^-	no γ
³⁶ Cl	310,000 y	0.714 β^-	
³⁸ Cl	37.3 m	1.1 to 4.8 β^-	1.6 to 2.2 γ
⁴⁰ K	4.5×10^8 y	1.3 to 1.9 β^-	1.46 γ
⁴² K	12.5 h	2.0 or 3.6 β^-	1.53 γ (18%)
⁴⁵ Ca	164 d	0.25 β^-	no γ
⁵² Mn	5.72 d	0.58 β^+ or E.C.	0.73 to 1.45 γ
⁵⁶ Mn	2.58 h	0.65 to 2.81 β^-	0.8 to 2.1 γ
⁵⁵ Fe	2.94 y	E.C.	
⁵⁹ Fe	45.1 d	0.27 or 0.46 β^-	0.2 to 1.1 γ
⁶⁴ Cu	12.8 h	0.66 β^+ or 0.57 β^- or E.C.	some γ
⁷⁶ As	26.5 h	0.4 to 3.0 β^-	0.6 to 2.0 γ
⁷⁷ As	38.7 h	0.2 to 0.7 β^-	some γ
⁸⁰ Br	4.6 h	I.T. some β	some γ
⁸² Br	35.9 h	0.44 β^-	0.5 to 1.5 γ
⁸⁶ Rb	18.6 d	0.7 or 1.8 β^-	some γ
⁸⁹ Sr	50 d	1.46 β^-	
⁹⁰ Sr	28 y	0.61 β^-	no γ
¹³¹ I	8.0 d	0.25 to 0.82 β^-	0.08 to 0.72 γ

E.C. = orbital electron capture by nucleus

I.T. = isomeric transition

 In cases where positrons are emitted, annihilation radiation (0.5 MeV γ) is always observed. For further details of the more complicated decay schemes see A.E.R.E. Catalogue No. 4 of Radioactive Materials, or Kamen¹, etc.

atoms that they approach sufficiently close, causing them to dissociate into an ion pair—a heavy positive ion and a negatively charged electron. Each such event or 'ionization' robs the moving particle of some of its energy, so that eventually it comes to rest. In air, roughly 30 electron volts are needed for the formation of a single ion pair, so that a β -particle emitted by ³H will be brought to a standstill in air after producing about 600 ionizations while one from ⁴²K will give rise to no less than 120,000 ion pairs. In the energy range with which we are concerned here the efficiency of the ionization process is inversely proportional to the square of the velocity of the β -particle, so that most of the ionization produced by a β -particle occurs near the end of its path, and the ionization per unit length of path is initially much less for an energetic particle than for a slowly moving one. The total distance traversed by a β -particle before it comes to a stop, that is to say its penetrating power, increases with its initial energy. In solid matter β -particles will not, of course, travel as far as in a gas, since interactions with

INTERACTIONS OF RADIATION WITH MATTER

atomic electrons will be spaced much closer together. If, however, the range of penetration is calculated in units of weight per unit area (i.e. distance \times density) it is found to be nearly independent of the nature of the absorbing material, provided that the absorber is not a very heavy element. Some idea of the penetrating power of homogeneous β -particles of different energies may be obtained from *Table 2*. The density of aluminium is 2,600

TABLE 2
Extrapolated Ranges for β -particles of Various Energies (from Kamen¹)

Energy in MeV	Range in mg/cm ² of aluminium
0.053	6
0.13	20
0.22	48
0.31	81
0.600	213
1.022	426
1.80	812

(By courtesy of Academic Press)

mg/cm³, so that a 1 MeV β -particle would be brought to rest by about 1.6 mm of aluminium or 4.2 mm of water; in air its range would be about 350 cm.

α -Particles, by virtue of their high charge and low velocity, cause a relatively intense ionization of the matter through which they pass, and their total range is correspondingly small. Thus the range of a 1 MeV α -particle in air is only about 5 mm. This results in some differences between methods for detecting α - and β -particles, but these need not concern us here since those isotopes which emit α -particles are not of major importance in biological research.

When γ -rays pass through matter they do not cause ionizations directly as the charged particles do, but instead they interact in one of three main ways with the electrons in the absorbing material. For γ -rays of relatively low energy the principal mechanism of interaction is photoelectric absorption. This consists in an encounter between the γ -ray photon and an extranuclear (orbital) electron, as a result of which all the γ -ray energy is transferred to the electron, ejecting it from its parent atom with a kinetic energy equal to the difference between the γ -ray energy and the energy with which it was bound to the atom. The electron subsequently loses the energy by causing ionizations in the manner just described for β -particles. The probability that photoelectric absorption will occur increases sharply with the atomic number of the absorber, being proportional to Z^4 . The second mechanism of energy transfer is termed Compton scattering, this being the dominant process in γ -ray absorption for energies between about 0.5 and 5 MeV. It consists of the transference of some of the γ -ray energy to an electron which may be considered as free (the quantity of energy transferred is generally large compared with the atomic binding energy, so that whether the electron is initially free or extranuclear makes little difference). After collision, the γ -ray photon

ASSAY OF RADIOACTIVITY

recoils with reduced energy in one direction, and the electron is ejected in another; the interaction behaves as a classical two-body collision, the angle through which the photon is scattered being that required for proper conservation of energy and momentum. A given γ -ray photon may take part in several collisions of this sort, losing some energy on each occasion, thus being degraded to a longer wavelength until in the end it is completely absorbed by the photoelectric effect. As before, the electrons involved in Compton scattering dissipate their energy in secondary ionizations. The third mechanism of energy transfer, known as pair production, is most important for very energetic γ -rays. If the photon has an energy greater than that equivalent to twice the rest mass of an electron (just over 1 MeV) it is possible for it to disappear within the Coulomb field of an atomic nucleus, and for an electron pair consisting of one positron and one electron to be formed instead. These receive the excess energy as kinetic energy, not necessarily in equal shares, the nucleus playing a part in conserving momentum. The probability that pair production will occur increases with γ -ray energy, and is proportional to the square of the atomic number of the absorber.

The range of penetration of γ -rays into an absorbing material depends on the sum of the probabilities of these three processes of energy transfer, and the overall absorption coefficient varies in a complicated way with γ -ray energy. In general, γ -rays lose much less of their energy per unit path length than do β -particles, so that their penetrating power is much greater; also, in contrast to β -particles, they are absorbed more effectively in materials of high atomic number, so that in air or water they can traverse great distances. Since the absorption by any given material increases exponentially with absorber thickness, γ -rays cannot strictly be said to be completely absorbed, but are attenuated in passing through matter. Their penetrating power is best expressed in terms of half-thickness values, that is the thickness of absorber required to reduce the energy of a beam of γ -rays to half the incident value. For 1 MeV γ -radiation the half-thickness of an aluminium screen is about 12 g/cm², or about 45 mm; for comparison, the half-thickness for 1 MeV β -particles is only 0.3 mm Al.

DETECTION OF IONIZING RADIATION

All methods for the quantitative determination of radioactivity depend on detection of the ionizations produced when α -, β - or γ -rays pass through solid, liquid or gaseous media. The techniques available may be divided into three main classes as follows:

Radioautography

The path of ionizing radiation through a photographic film is marked by a blackening of the grains in the emulsion similar to that produced by visible light. This effect was first exploited over 60 years ago to demonstrate the presence of radioactivity. It has not often been used for precise quantitative measurements, but provides the most convenient basis for visualizing the distribution of radioactivity in tissue sections or small whole organisms. Developments in recent years have improved the spatial discrimination of

BASIC PHENOMENA IN GAS CHAMBERS

radioautography, and under favourable conditions it is possible to achieve a resolving power of the order of a few microns. Since this technique does not directly involve the use of electronic apparatus it will not be discussed in detail here, and the reader is referred for further information to the general texts already mentioned and to reviews by Doniach, Howard and Pelc⁸, and Boyd⁹.

Scintillation counting

Sir William Crookes observed in 1903 that the impact of α -particles on a screen of zinc sulphide resulted in the emission of flashes of light. This phenomenon depends on the formation of 'excited' atoms by the displacement of electrons from atoms in a crystal lattice, and the subsequent emission of radiation in the visible spectrum (and beyond it) when the excited atoms revert to the unexcited ('ground') state. It formed the basis for the visual counting techniques used by Rutherford and others in the early researches on nuclear physics, but visual scintillation counting had severe limitations, and was later abandoned in favour of methods depending on the use of gas chambers (see below). In the last few years the simultaneous development of improved methods for converting brief flashes of light into electrical pulses, and of more convenient 'phosphors' (the name given to solids or liquids in which ionizing radiation can be converted to visible or ultraviolet light), has brought scintillation counting back into use, particularly for γ -radiation. The apparatus needed for scintillation counting is discussed in a later section of this chapter.

Gas chamber counting

The majority of techniques used in the assay of radioactivity depend on the ionization arising from the passage of radiation through gases. It is possible either to observe the integrated effects of a number of ionizing particles or quanta of radiation, as in *ionization chambers*, or to detect the emission of single particles, as in *proportional counters* and *Geiger-Müller counters*. The relationship between these three types of gas chamber and their special features are considered in the following sections.

BASIC PHENOMENA IN GAS CHAMBERS

As has already been seen, when radiation interacts with the atoms in a gas, pairs of ions are formed, each consisting of one negative electron and one heavy positive ion. The charged α - and β -particles ionize directly, while γ -radiation causes ionization indirectly by setting in motion secondary charged particles. If the gas in which the ionization is occurring is contained in a chamber provided with electrodes maintained at an appropriate potential difference relative to one another, the electrons will tend to move towards the anode and the positive ions will move towards the cathode. The amount of charge collected varies with the applied potential in a characteristic way, as shown in *Figure 31.1*. Here the charge collected at the anode is plotted logarithmically against the anode voltage for two different situations: (I) for an initial ionization consisting of only a few ion pairs, as might be produced by the passage of a β - or γ -ray through the sensitive volume between the

ASSAY OF RADIOACTIVITY

electrodes, (2) for an initial ionization involving a much larger number of ion pairs, as might be formed by an α -particle. Starting from zero potential, there is first a region (*A*) in which the collected charge increases rapidly with voltage. At these relatively low potentials (up to about 50 V) only some of the electrons are collected, the remainder being lost by re-combination with positive ions before they have time to reach the anode. The fraction collected rises until a point is reached where the members of every ion pair are separated fast enough to prevent their re-combination. Above this potential there is

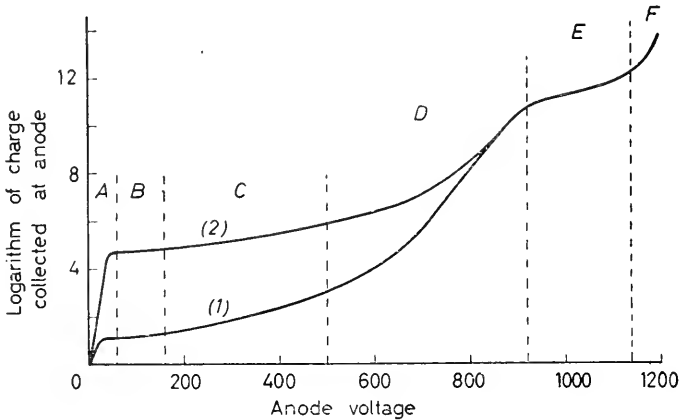


Figure 31.1 The variation with applied voltage of the charge collected at the anode of a gas chamber—for explanation of symbols see text

a plateau region (*B*) where collection is complete, and where the current is directly proportional to the number of ion pairs liberated within the sensitive volume in unit time. In this 'saturation' region the charge is nearly independent of the applied voltage, but varies considerably with the nature of the incident radiation, as indicated by the difference between curves 1 and 2. If the electrode potential is further increased some of the electrons acquire sufficient kinetic energy while being accelerated towards the anode to be able to produce additional ion pairs in their turn, this process being known as gas multiplication. In region (*C*), the zone in which proportional counters are operated, the multiplication factor increases somewhat with voltage, but at a given voltage each primary ion pair is multiplied to the same extent, so that curves 1 and 2 rise in parallel, and the amount of charge collected is still dependent on the nature of the incident radiation. There follows a transition region (*D*) in which the multiplication factor rises still further, and in which the number of electrons reaching the anode ceases to be proportional to the number liberated in the initial ionization. At a sufficiently high voltage the Geiger-Müller region (*E*) is reached. Here curves 1 and 2 coincide, a single ion pair causing an avalanche of electrons to reach the anode whose magnitude depends only on the characteristics of the chamber. Still higher voltages result in a continuous discharge (*F*).

Figure 31.1 is purely diagrammatic, and the exact lengths of the various regions and the slopes of the curves depend in a complicated way on the

IONIZATION CHAMBERS

type of gas used to fill the chamber, the pressure, and the dimensions of the system. The effects of these factors are discussed by Rossi and Staub¹⁰ and Wilkinson¹¹.

IONIZATION CHAMBERS

Ionization chambers are operated in the plateau region (B) of *Figure 31.1*, where the total charge collected is ideally equal to the total ionization within the sensitive volume, the potential being high enough to achieve saturation

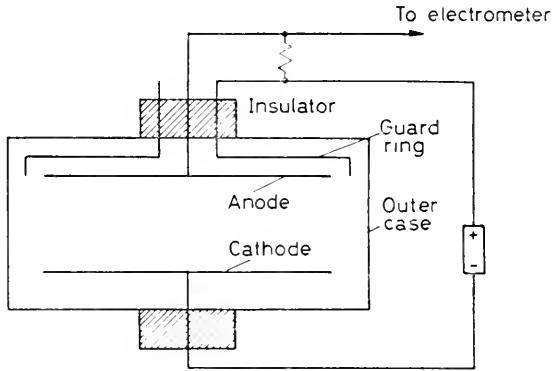


Figure 31.2 Essential features of a parallel-plate ionization chamber

but not so high that there is any gas multiplication. *Figure 31.2* shows the essential features of a parallel-plate ionization chamber, the only one needing further description being the guard electrode, maintained at the same potential as the collecting electrode and serving both to define precisely the sensitive volume of the chamber and to prevent leakage of current across the insulator on which the collecting electrode is mounted. In order to measure the charge collected the chamber must be connected to some type of electrometer (i.e. to a high impedance device for measuring small potential

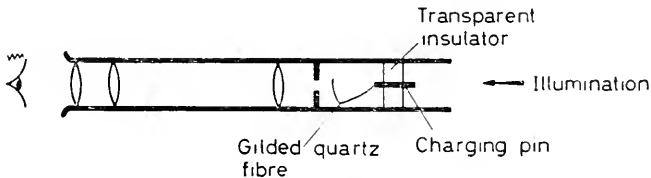


Figure 31.3 Pocket quartz-fibre dosimeter

changes), the most satisfactory being the vibrating reed electrometer, in which a form of variable capacitor serves to convert the steady potential change across the capacitance of the collecting electrode and the rest of the system into an oscillating voltage which can be amplified by a conventional a.c. amplifier. An alternative method of measuring the charge is to combine the chamber with a quartz-fibre electroscop of the type developed by Lauritzen. In this instrument the metal-coated quartz fibre (see *Figure 31.3*) is initially displaced by charging it to a potential of the order of +100 V

ASSAY OF RADIOACTIVITY

relative to the metal case; when negatively charged electrons are collected the fibre moves back by an amount which is, for small displacements, linearly related to the total charge, and hence to the total ionization that has occurred within the sensitive volume. The movements of the fibre can be observed directly with a lens system, against a scale calibrated in standard units of radiation (the roentgen).

Since ionization chambers do not take advantage of the magnification that can be obtained by gas multiplication, their sensitivity is low. If, for example, the total capacitance of the collecting electrode system is $10 \mu\mu\text{F}$, the potential change on collection of a single electron is only about $0.016 \mu\text{V}$, which is not measurable. In contrast, therefore, to proportional and Geiger-Müller counters, they are inconvenient for the detection of individual ionizing particles, but will measure the integrated ionization resulting from larger amounts of radiation. This limits their application in biological tracer experiments, where maximum sensitivity is usually desirable, and often essential. Used in conjunction with a vibrating reed electrometer, ionization chambers have been applied successfully by some workers to the determination of ^{14}C , introduced directly into the chamber as CO_2 , but their principal value is for monitoring large sources of radiation. A number of radiation monitors using ionization chambers are available commercially, as are pocket quartz-fibre dosimeters, which are especially valuable for measuring the total doses of radiation received by an individual in the course of handling radioactive materials.

FORMATION OF THE AVALANCHE IN GEMER-MÜLLER COUNTERS

The basic construction of two commonly used types of Geiger tube is illustrated in *Figure 31.4*. In both of them the anode is a thin wire coaxial with a cylindrical cathode; normally they are used with the cathode at earth potential, and the anode at a suitable positive potential. They differ in the thickness of wall through which the incident radiation has to penetrate before it reaches the sensitive volume. The 'end-window' tube has a relatively thick metal cathode, but its end window, usually of mica, can weigh as little as 1.0 mg/cm^2 , thus permitting the entry of even the weakest β -radiation (excepting that of ^3H) without excessive absorption by the window. The 'thin-walled' tube is designed for the measurement of rather stronger β - and γ -radiation, having a thin glass wall weighing about 30 mg/cm^2 ; the cathode, immediately inside the glass wall, may take the form of a layer of carbon, an open spiral of stainless steel wire, or (particularly for soft γ -radiation) of a thin sheet of lead.

It has already been mentioned that in the Geiger-Müller region (*E*) of *Figure 31.1* the formation of a single ion pair within the sensitive volume of the chamber leads to the discharge of a very large number of electrons, the so-called 'Townsend avalanche', at the anode. The mechanism of this process must now be considered in slightly greater detail. Immediately after an ionizing particle (either emitted directly by a β -emitter or produced indirectly by the passage of a γ -ray) has formed a few ion pairs in the gas (usually argon or neon) filling the chamber, the electrons begin to move towards the anode;

FORMATION OF THE AVALANCHE IN GEIGER-MÜLLER COUNTERS

their motion is rapid compared with that of the much heavier positive ions liberated at the same time. As the electrons approach the thin wire anode they are increasingly accelerated, since the voltage gradient is greatest near the wire. Soon they are travelling fast enough to cause further ionizations by collisions with the gas molecules, giving rise to secondary electrons, which in their turn are attracted to the anode. The secondary electrons may then breed tertiary electrons, and so on. The consequence of this cumulative process is that a shower of over 10^8 electrons reaches the anode. However, the discharge of electrons at the anode ceases after a period of the order of

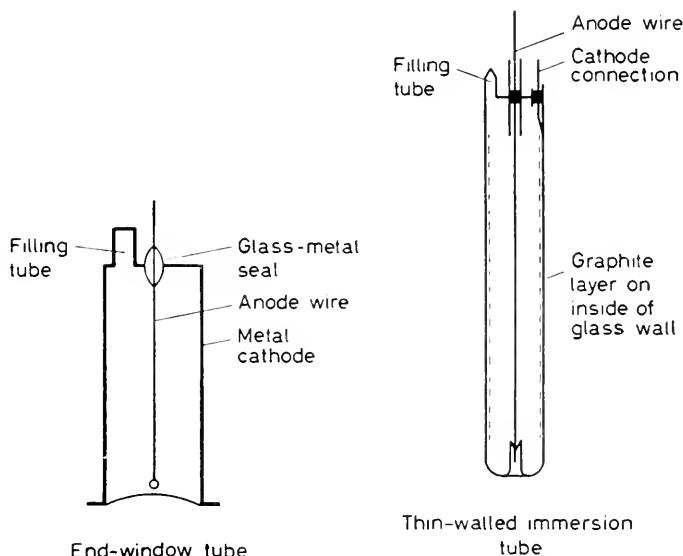


Figure 31.4 Basic features of two types of Geiger tube

0.1 μsec , because the slow-moving positive ions build up a virtually stationary space charge which neutralizes the effective voltage gradient in the vicinity of the anode. Before the tube can discharge again the positive ions have to be sufficiently dispersed towards the cathode to restore the voltage gradient at the anode; this occurs relatively slowly, and may take 100 μsec or more. The period during which the tube is thus rendered inoperative is called its 'dead time'. It is followed by a rather longer interval (while the positive ions are still being collected at the cathode) during which the tube gives an output pulse of reduced size if triggered again.

Quenching mechanisms

When the avalanche of electrons reaches the anode, photons of electromagnetic radiation are produced. Unless precautionary steps are taken, these are now capable of initiating a fresh discharge by ejecting electrons from the material of the cathode, thus setting up a state of continuous discharge. The limiting condition for a continuous discharge is that each avalanche should give rise to just one photoelectron at the cathode, i.e. that the product of the gas multiplication factor and the photoelectric efficiency

OUTPUT CHARACTERISTICS OF GEIGER TUBES

into fragments, rather than by photon emission. The chances of initiating a spurious discharge are therefore substantially reduced.

Alcohol as a quenching agent has the disadvantage that its dissociation products cannot reconstitute the original molecules, so that a small quantity of alcohol is used up during each discharge. This limits the useful life of the tube, and means that it may be irretrievably damaged if it is allowed to go into a continuous discharge (which may occur despite the presence of alcohol if too high an anode voltage is applied) for any length of time. In halogen-quenched tubes the halogen atoms formed on dissociation can subsequently re-combine, so that such tubes have a very long life and are less likely to be harmed if they accidentally pass into a continuous discharge. Halogens have the additional advantage that they can be used to quench low-voltage Geiger tubes filled with argon-neon mixtures; such tubes have operating potentials around 500 V whereas conventional argon-alcohol tubes need about 1,200 V. The main disadvantages of halogen-quenched tubes are that they are at present more expensive, and that they have somewhat less than 100 per cent efficiency for recording β -particles within their sensitive volume. However, their greater cost is offset by their longer life, and except for making absolute measurements or coincidence experiments it does not greatly matter if a small (but, of course, fixed) proportion of counts is missed.

OUTPUT CHARACTERISTICS OF GEIGER TUBES

The discharge of electrons at the anode and positive ions at the cathode causes a transient flow of current through the Geiger tube, and the potential of the centre wire drops temporarily. This fall in potential constitutes the output pulse, which is then fed via a probe unit to a device for measuring the rate of arrival of pulses, i.e. to a scaling circuit or ratemeter. Much the greater part of the output pulse is contributed by the positive ions, since they have to move through most of the applied potential gradient before reaching the cathode, whereas the electrons only move a short distance to the anode. The size of the output pulse depends on the total amount of charge on all the positive ions (which increases with the length of the tube and with the applied potential), on the total capacity (C) between anode and cathode, to which it is inversely proportional, and on the recovery time constant RC , where R is the parallel resistance of the EHT feed resistor and the input resistance of the probe unit (see *Figures 31.5* and *31.6*).

For a given value of C , maximum pulse amplitude requires a value of RC large compared with the total period (200–500 μsec) needed for complete collection of the positive ions; but if R is made large enough to achieve this, the potential on the centre wire returns excessively slowly to its resting level after each pulse, which may be disadvantageous if fast counting rates are to be dealt with. In practice it is usual to compromise by making RC about 50 μsec . This gives pulses of about half the maximum size, generally of the order of a few volts. In order to avoid loading the tube with a long length of cable, thus increasing C and reducing pulse size, the probe unit is mounted close to the tube. There is something to be said (see below) in favour of using the type of quench probe shown in *Figure 31.5*, even with a self-quenching Geiger tube; but with such tubes it is not essential to provide a

ASSAY OF RADIOACTIVITY

quenching pulse, and the straightforward probe circuit illustrated in *Figure 31.6*, which either inverts and slightly amplifies the pulses or can be used as a cathode follower, has the merits of greater simplicity and reliability.

Various types of scaling circuit are described in Chapter 41. Until recently, most workers used hard-valve scalers incorporating several scales of eight or,

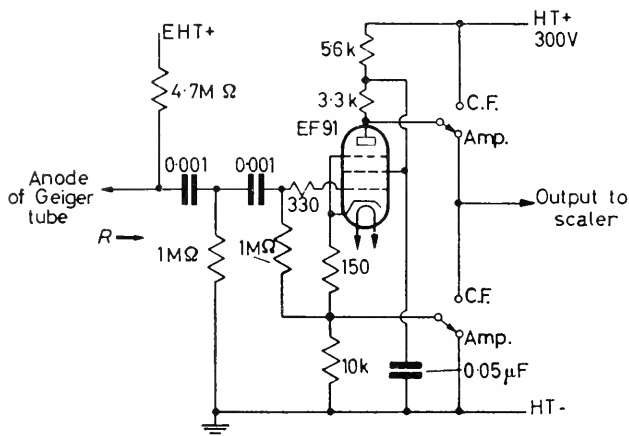


Figure 31.6 A simple probe circuit (A.E.R.E. Probe type 200)

more conveniently, of ten, arranged in cascade, driving a mechanical register: but the development of Dekatron counting tubes seems likely to supplant the hard-valve circuits, except for the first stages of scalers designed to operate at very high counting rates. The sensitivity of the scaler, taking into account some amplification of the pulses by the probe unit, is normally such that a

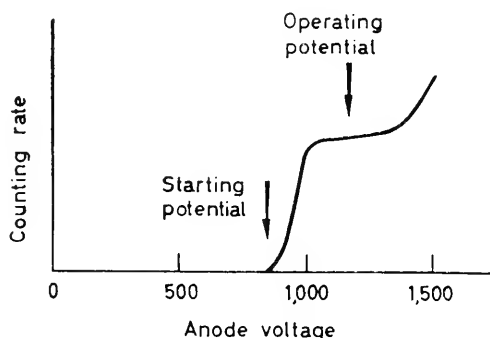


Figure 31.7 Characteristic curve for an argon-alcohol filled Geiger tube

few pulses begin to be recorded when the applied potential is just below the Geiger-Müller region (*Figure 31.1*). The voltage at which counting begins is called the starting potential, and the curve obtained by plotting the counting rate for a fixed source against applied potential is the characteristic curve of the counter. A typical curve for an argon-alcohol filled tube is shown in *Figure 31.7*.

COUNTING ERRORS: BACKGROUND AND DEAD TIME

At the starting potential the pulses produced by the Geiger tube are somewhat variable in size since this potential is in the transition region of *Figure 31.1*, and the amplitude of the output pulses is still governed to some extent by the number of ion pairs formed by the incident radiation. Under these conditions only the largest pulses are able to trigger the scaler. As the applied potential is raised, more and more of the pulses exceed the threshold of the scaling circuit, and the counting rate increases rapidly, until when the Geiger-Müller region is reached the pulses are uniform in size and are all large enough to be recorded. A further rise in potential beyond this point makes the pulses even larger but has little effect on the counting rate, so that the characteristic curve shows a flat plateau. The normal operating voltage of a Geiger tube is in the centre of this plateau, where small variations in the applied potential cause the minimum error in the recorded count. In a well designed tube the length of the plateau should be about 200 V, so that some latitude is permissible in choosing the operating potential. At the upper end of the plateau the counting rate begins to rise steeply again, largely because at excessive operating potentials the quenching vapour ceases to be able to suppress all the spurious counts; the tube is apt to be damaged if it is operated for any appreciable period at such a potential.

The slope of the plateau is not zero, although in a good tube it should not be greater than about 0.1 per cent per V. The finite slope arises partly through an increase in the sensitive volume with operating potential, and partly through production of a small but increasing proportion of spurious pulses. The change in sensitive volume is an end-effect which is most marked in small tubes, and which can never be avoided altogether. The effect of spurious counts can, however, be virtually eliminated by using a quench probe, which enables an appreciably flatter plateau to be attained; but in general it is unnecessary to take great pains to flatten the plateau, since it is not difficult to design an EHT supply unit (see Chapter 37) whose output is sufficiently well stabilized for errors arising from shifts up and down the characteristic curve to be small compared with those from other sources.

When Geiger tubes have been in use for some time there may be a tendency for the plateau to shorten or alter its position. In order to keep a check on the condition of a tube it is therefore highly desirable to make routine measurements of the starting potential and plateau characteristics.

COUNTING ERRORS: BACKGROUND AND DEAD TIME

The quantity ultimately to be measured in almost all tracer experiments is, in effect, the number of labelled atoms in a given sample. The observation actually made consists in counting the number of atoms which, over a known period, disintegrate to give detectable radiation. It is convenient at this point to consider briefly the various sources of error that have to be taken into account when making such observations.

The rate at which a given species of radioactive isotope decays is a property of its nuclear constitution, and for all practical purposes is independent of its physical and chemical state. The law of radioactive decay can be written as

$$dN/dt = -\lambda N$$

ASSAY OF RADIOACTIVITY

where N is the number of atoms present at time t , λ is the decay constant for the particular isotope, and dN/dt is the rate of disintegration of the nuclei. This expression embodies the observation that the chances of any nucleus disintegrating within a given period are constant, no matter how long it has already existed. It leads, of course, to the familiar exponential expression

$$N = N_0 e^{-\lambda t}$$

where N_0 is the number of atoms at zero time, and it will be seen that λ is related to the half-life of the isotope (see *Table 1*) by

$$t_{1/2} = 0.693/\lambda$$

Now since the individual disintegrations are not causally related to one another, they occur not at regular intervals but at random. This has the important consequence that if a sample is counted under fixed conditions several times within a period which is short compared with its half-life, the results will not be the same on each occasion but will vary according to the Poisson distribution law for random events. There is not space here to discuss the statistics of counting in detail, but some of the practical repercussions of the random fluctuations in counting rate that thus inevitably occur in all radioactivity measurements must next be considered.

In the first place, in order to have reasonable confidence that an experimentally determined counting rate can be relied upon to be close to the true counting rate for a sample, it is necessary to record a fairly large number of counts from the sample. The more counts are recorded, the greater will the reliability of the result be. For a Poisson distribution, the standard error (S.E.) of a single determination of a counting rate is the square root of the total number of counts recorded (this assumes, as is normally the case, that any error in the timing is so small that it can be neglected). The S.E. is a standard way of expressing the probable scatter of the results; rather roughly, there is a 1 in 20 chance that the value to which it is attached differs from the true value by more than twice the S.E. To take a specific example: if only 100 counts were recorded from a sample in x minutes, the counting rate would be $100/x$ and its S.E. would be $10/x$ —that is to say the S.E. would be ± 10 per cent of the recorded result, which could not be regarded as a very reliable estimate of the true counting rate. If 10,000 counts were recorded in y minutes, then the counting rate would be $10,000/y$ and its S.E. would be $100/y$, which is only ± 1 per cent of the recorded result. In most biological experiments this would be an acceptable answer, since it is generally difficult to reduce other sampling errors much below 1 per cent; but the counting rate could be determined to within still closer limits by prolonging the counting period so as to record an even larger number of counts. Conversely, it may often be unnecessary to achieve an S.E. as low as ± 1 per cent, and shorter counting periods will then yield sufficiently accurate results.

What has just been said is strictly true only for rather active samples, since for weak samples a further complication is introduced by the *background count* of the Geiger tube. Most arrangements for recording ionizing radiation will give a slow counting rate even when there is no radioactive source close to them because of cosmic radiation and traces of radioactivity

which are almost always to be found in the floor and walls of a room. Geiger tubes are also sensitive to light, and the background will be further increased if they are not operated in complete darkness. In order to reduce the background to a minimum, both the Geiger tube and the sample are, if possible, enclosed within a thick lead shield or castle, but even with 3 in. of lead on all sides a Geiger tube of conventional size (say 1 in. in diameter and 3 in. long) gives a background of the order of 10 counts/minute. The background tends to be proportional to the sensitive volume of the tube, so that larger tubes will inevitably give still higher background counts no matter how well shielded they are. The background can normally be treated as constant, but it is subject to the same statistical fluctuations in rate as the count given by a radioactive sample. When dealing with weak samples it is therefore necessary to determine the background in the absence of the sample, and to subtract it from the total counting rate obtained for sample plus background. The S.E. of the background determination can often be made small enough to be neglected, by taking very long background counts, but the S.E. of the corrected counting rate for the sample is calculated from the square root of the total counts given by sample plus background, so that for samples giving counts not much greater than background, the S.E. of the final result is apt to be unsatisfactorily high in comparison with the result. Furthermore there are, in the limit, statistically significant variations in the background from time to time, arising from phenomena such as cosmic ray showers. It is, consequently, difficult to make reliable measurements of samples giving less than, say, half the background count. When confronted with samples as weak as this, little can be gained by greatly prolonging the counting periods, and it is necessary to resort to other means of improving the accuracy of the measurements. Thus it may be possible to re-design the experiments so as to increase the activity of the samples, or to improve the geometry of the counting arrangement (see below) so that a higher proportion of the ionizing particles emitted by the sample penetrates the sensitive volume of the Geiger tube. The effective background can also be reduced, at the cost of appreciable complications in the apparatus, by employing coincidence counting techniques.

In the other direction, when excessively active samples are to be counted a limitation is set by errors arising from the dead time of the Geiger tube and associated equipment. As has already been mentioned, once a Geiger tube has been triggered by an ionizing particle there is a period of the order of 100 μ sec during which a second particle entering the sensitive volume will not be detected. This dead time is irreducible, and may, in fact, be deliberately increased by operating the tube in conjunction with a quench probe set to give a quenching pulse which lasts appreciably longer than 100 μ sec. In view of the random spacing between disintegrations some particles are bound to arrive during the dead time, so that the apparatus always fails to record a certain proportion of them. This proportion is not fixed, but rises steeply with counting rate; thus for a dead time of 100 μ sec, less than 2 counts/min are missed at a true rate of 1,000 counts/min, but at 5,000 counts/min about 40 counts/min are lost—an error of nearly 1 per cent. Provided that the dead time (τ) is known, it is simple to calculate the correction that should be made to the recorded counting rate (N) in order to arrive at the true counting rate

ASSAY OF RADIOACTIVITY

(N_0), since it can be shown that the relationship between these quantities is

$$1/N_0 = 1/N - \tau$$

(if a table of reciprocals is available, this is probably the most convenient form of the relationship). However, complication may arise through variation in τ , since for the Geiger tube itself the dead time varies with operating voltage, being perhaps twice as great near the starting potential as it is at the upper end of the plateau. This difficulty can be overcome by determining τ under the standard operating conditions, by counting two samples, one of them a very active one, whose relative activities are known precisely beforehand (e.g. by a known dilution): but this determination will have to be repeated for each individual Geiger tube, and it may be preferable to work with a longer, but strictly reproducible, dead time determined by a quench probe. In this case the value of τ can be measured with two samples as before, measured with a multiple pulse generator, or (which is most convenient of all) pre-set by the makers of the quench probe.

Whether or not a quench probe is employed it is advisable not to work at such high counting rates that large dead-time corrections are called for. Counting rates can usually be reduced to manageable proportions by dividing up the samples or by placing them further from the Geiger tube.

For the sake of completeness, other obvious sources of error in radioactivity measurements are as follows:

Decay of the isotope—It often happens that measurements are spread over a period of the same order as the half-life of the isotope. In this case it is necessary to note the time at which each count is made, and subsequently to correct all results to the same zero time, using the value for the decay constant of the isotope. It is generally advisable to check for each isotope sample that the rate of decay is what it should be; this provides a guard against contamination by unwanted isotopes. It may also be noted that differences in half-life can be used as a means of distinguishing between the components of mixtures of isotopes in multiple labelling experiments.

Positioning errors—Radioactive samples emit their radiation unselectively in all directions, but it is only the radiation which penetrates into the sensitive volume of the Geiger tube that is recorded. The counting rate obtained for a given sample is therefore roughly proportional to the solid angle subtended by the counter at the sample, and may be considerably altered by small movements of the one relative to the other. In making absolute measurements of the radioactivity of a sample, determination of the effective solid angle within which counts are recorded is a matter of some difficulty, but this problem does not arise in most tracer experiments since the measurements usually made are relative ones, the activity of an unknown being compared with that of a standard sample. It will be clear, however, that positioning errors will arise unless care is taken to see that all samples and standards are counted in exactly the same position. Alternatively, if the total solid angle for counting can be increased to 4π or near it (e.g. by placing the sample within the sensitive volume of the counter tube or by surrounding it with a ring of tubes), positioning errors will become less important.

TYPES OF GEIGER TUBE

Absorption errors—Any radiation that is absorbed by the window or wall of the Geiger tube, or within the sample itself, will not be recorded, thus cutting down the counting rate. Self-absorption by the upper layer of radiation emitted by the lower layer of a solid sample is a particularly dangerous source of error when working with soft β -emitters. Again, the usual way of overcoming this source of error is not to correct for it, but to ensure that there is the same degree of self-absorption in both unknown samples and in standards. This means that in counting solid samples, unknowns and standards must occupy the same area (if they do not, there may also be positioning errors) and have the same total weight of material per unit area, while in using liquid counters, densities must be kept the same in unknowns and standards. When attempting to balance out errors in this fashion it is, of course, desirable to keep all absorption losses as low as possible, using the most appropriate type of counter for any given isotope.

Scattering errors—To radiation travelling directly from the sample into the sensitive volume of the Geiger tube is added a certain amount of radiation scattered from the walls of the enclosure in which tube and sample are contained, from the support on which the sample is mounted, and from within the sample itself. It is possible to correct for such scattering effects, but simpler once more to see that they are the same for unknowns and standards. It should be noted that back-scattering from the sample mount may make an appreciable contribution to the total count.

Timing of counts

For many purposes it is sufficient to time the counts of isotope samples by hand, using a stop-watch. In handling large numbers of samples it is, however, more convenient to use some form of automatic timer. Timers incorporating a standard time source and scaling circuits have been devised (and are available commercially) to give a choice of two methods of controlling a scaler: (1) to determine the time taken for a pre-set number of pulses to be recorded, (2) to determine the number of pulses recorded during a pre-set time interval. The advantage of method (1) is that it pre-determines the standard error of each count. A timer of this sort is complex and expensive, and the writer has always used a simple automatic timer constructed from a cheap clock and a relay (see Cook and Keynes¹²), which over a number of years has proved entirely reliable.

TYPES OF GEIGER TUBE

End-window counters

Probably the most commonly employed type of Geiger tube is the end-window counter illustrated in *Figure 31.4*. End-window tubes are available from several makers, e.g. Mullard (type MX123), G.E.C. (types GM4 and EHM2S), 20th Century Electronics (type EW3H), and others. Most of these tubes have very thin mica end-windows, making them suitable for work with soft β -emitters like ^{14}C and ^{35}S ; but tubes with duralumin or glass end-windows are also manufactured, these being more robust but also somewhat thicker. Some are high-voltage tubes quenched with alcohol

vapour, others are low-voltage halogen-quenched tubes; the relative merits of the two classes have already been discussed.

End-window tubes are normally used in lead castles fitted with shelves for supporting at various heights beneath the window a tray carrying the radioactive sample in solid form. The samples are usually mounted on thin metal planchettes. With this arrangement the top shelf brings the sample close enough to the window for up to 30 per cent of the radiation emitted to reach the sensitive volume. The lower shelves are useful if the samples are giving excessively high counting rates or if it is necessary to insert a filter between the sample and the tube in order to cut out weak radiation in experiments involving mixtures of different isotopes. This type of mounting can also be adapted for use with automatic sample changers, in which a succession of samples is brought under the tube by a mechanical device, and the counts are recorded automatically.

End-window tubes are also convenient for a variety of other purposes. Thus they may be used to scan paper strips on which labelled compounds have been separated by chromatography. They may be used for general monitoring. They may be applied externally to living animals in order to locate concentrations of radioisotopes within the animal (e.g. in experiments on circulation time, or the uptake of isotopes by different organs), or they may be used to measure changes in the amount of radioactivity in isolated living tissues, mounted in a thin-bottomed chamber in a stream of inactive fluid (an example of this type of application is described by Keynes¹³).

Liquid counters

It is sometimes more convenient to count samples in liquid than in solid form, and on such occasions a liquid counter may be employed. This incorporates the kind of thin-walled Geiger tube illustrated in *Figure 31.4*, immersed in a tube of such a diameter that the liquid occupies an annulus about 2 mm wide surrounding the sensitive volume. In the version designed by Veall¹⁴, which is available commercially from several of the makers listed in the previous section, the outer glass tube is fused with the Geiger tube to form a fixed skirt, this arrangement yielding a liquid counter with admirably reproducible geometry. These skirted tubes can be operated in two ways. About 9 ml of liquid is enough to cover the thin-walled part of the Geiger tube, and the addition of further active fluid above this level has little effect on the counting rate; if, therefore, it is used with about 10 ml of fluid, it measures the specific activity of the fluid. Alternatively, if a small volume (say 3 ml) of active fluid is poured into the counter and then progressively diluted by adding inactive fluid, it is found that the counting rate reaches a fairly flat maximum at a total volume around 6 ml. For determining the total activity in a sample whose precise volume need not be known if it is less than 6 ml, the tube can therefore be used by pipetting the sample into it, making up to a mark at 6 ml with inactive fluid, and mixing. Small variations in making up to the mark cause negligible errors.

Owing to absorption of radiation in the fluid, liquid counters are unsuitable for use with weak β -emitters like ^{14}C and ^{35}S , but give reasonable counting efficiencies for β -radiation stronger than 0.5 MeV, or for γ -radiation. They give an effective solid angle for counting of nearly 2π , and for a very strong

TYPES OF GEIGER TUBE

β -emitter like ^{42}K the counting rate for a given amount of the isotope is about the same for a skirted liquid counter used in the manner just described as it is for a conventional end-window counter. With weaker radiation the efficiency falls because of absorption losses, but is adequate for many of the isotopes of importance in biological work. In using liquid counters care must be taken to avoid errors from adsorption of the isotope on the glass surfaces, which in the author's experience may give trouble when working with an anion such as $^{32}\text{PO}_4$, but is not a serious problem with cations like ^{24}Na and ^{42}K . It is essential to wash out the tube thoroughly between

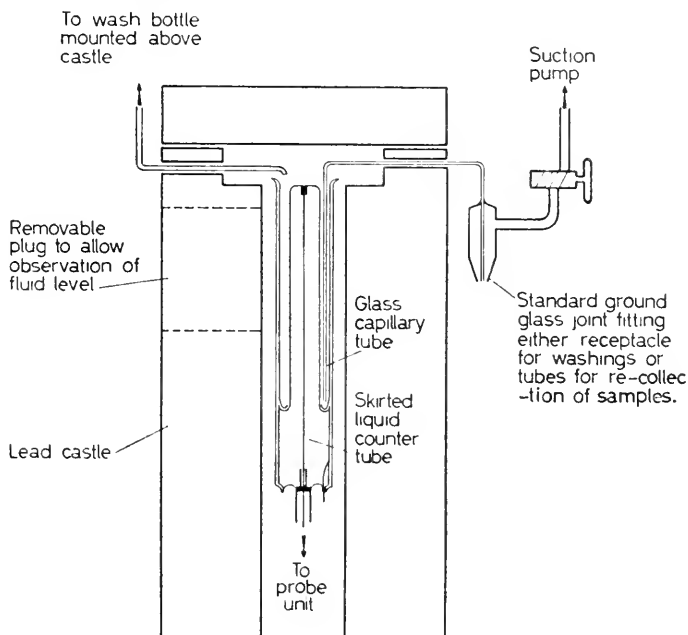


Figure 31.8 An arrangement for washing out liquid counter tubes without removing them from the castle

samples, either removing it from the lead castle for the purpose or else fitting the castle with an arrangement for washing *in situ*. The system shown in Figure 31.8, with a reservoir of wash-fluid mounted above the castle and a fine glass tube inserted into the counter in order to empty it by suction, proves very convenient to use. Two or three washes suffice to bring the counting rate down to background, even after a sample giving several thousand counts per minute.

Inside counters

For the weakest β -emitters, even a thin mica end-window can cause appreciable absorption losses, and some workers have therefore used demountable tubes, similar in shape to an end-window tube, in which solid samples can be counted inside rather than outside the gas space. With such an arrangement there are still losses from self-absorption in the sample,

and others have preferred to introduce the samples by adding them in gaseous form to the filling gas. This technique has the double advantage of completely avoiding absorption losses, and of achieving 4π geometry, so that it can be used for absolute as well as relative measurements of radioactivity; however it entails a somewhat complicated filling system. In order to exploit the undoubted advantages of counting isotopes like ^3H and ^{14}C in the form of gases, it is probably better to employ a proportional counter operating at atmospheric pressure, as discussed below.

γ -ray counters

As was pointed out earlier, γ -rays have only a low probability of producing an ionization in the gas space of a Geiger tube, so that the counting efficiency for γ -radiation is apt to be very small. The secondary electrons produced by photoelectric absorption and Compton scattering in the wall of the tube can, however, be detected satisfactorily, and the overall γ -counting efficiency of a Geiger tube can therefore be somewhat increased by building it with fairly thick walls and by using for these a material of high atomic number (e.g. lead). Even with this refinement the maximum efficiency of Geiger tubes for γ -ray detection is not more than 1 or 2 per cent. For monitoring γ -radiation, Geiger counters have the important merit of simplicity, but for quantitative measurements they have now been supplanted by the scintillation counters described below, whose efficiency is very much higher.

PROPORTIONAL COUNTERS

In recent years proportional counters—that is to say gas chambers operating in region (C) of *Figure 31.1*—have come into increasing use for the assay of radioisotopes. Their main advantages are that the dead time is very much shorter than in a Geiger tube, enabling them to be used at much higher counting rates, and that they are more stable in operation than Geiger tubes, being less affected by such factors as changes in temperature. Their most obvious disadvantage is that they produce relatively small pulses of current, and their output therefore has to be fed to a high-gain linear pulse amplifier in order to make the pulses large enough to trigger a scaler. They also necessitate a rather well stabilized EHT power supply, and, for some applications, elaborate discriminator circuits which select only pulses within a given size range. However, for most purposes it is sufficient to use a simple circuit which merely rejects all pulses below a certain size (i.e. to provide a threshold control for the scaler), and provision of a suitable amplifier is not difficult. *Figure 31.9* shows a circuit with a good enough frequency response to provide undistorted and linear amplification of the μsec pulses produced by proportional counter tubes.

It is possible to use a conventional end-window tube in the proportional region of its characteristic, but most proportional counters are of the 'flow' type, filled with continuously flowing gas at atmospheric pressure. Almost any gas can be used, but methane is the usual choice since it has some quenching action and allows a greater gas multiplication factor to be employed. Samples may be placed beneath a thin side-window, or introduced inside the counter with a sliding shelf arrangement. For the assay of tritium,

SCINTILLATION COUNTERS

which presents particular difficulty because of its exceedingly weak radiation, the isotope may be added to the filling gas either as methane or as molecular hydrogen. ^{14}C may similarly be counted as carbon dioxide or acetylene.

One of the respects in which proportional counters differ from Geiger counters is, of course, that the output pulses are proportional in size to the energy of the β -particles that gave rise to them. By suitably analysing the pulses with the help of discriminator circuits it is therefore possible both to sort out the radiations emitted by mixtures of isotopes in double-labelling

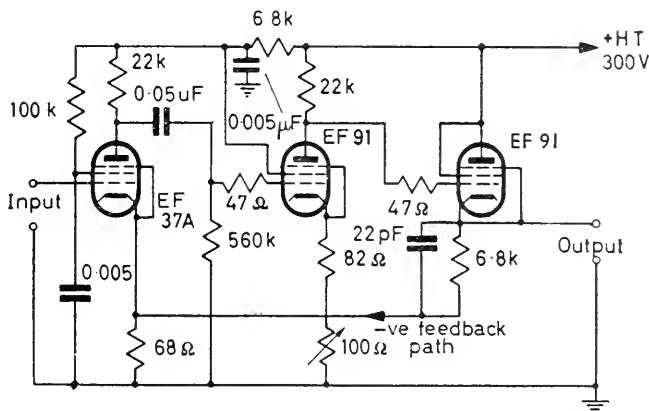


Figure 31.9 The basic elements of a negative feedback amplifier with a gain of the order of 100. The main section of A.E.R.E. Pulse Amplifier Type 1008, designed for use with proportional counter tubes, consists of two circuits of this type in cascade, separated by an attenuator

experiments, and to achieve very low background counts by rejecting all pulses outside the range characteristic of the isotope being used.

SCINTILLATION COUNTERS

As has already been mentioned, scintillation counting was first used over 50 years ago for the determination of α -particles, but as soon as gas chambers had been developed the technique fell into disuse. It is now coming into steadily increasing favour once more, since it not only provides a means for detecting γ -rays with much higher efficiency (up to about 50 per cent) than any gas chamber method, but has also been applied successfully to the determination of low-energy β -particles. The initial re-introduction of scintillation counting resulted from the availability of highly sensitive photomultiplier tubes (see Chapter 28) capable of converting the minute flashes of light produced by γ -rays in suitable crystals into electric pulses which could be counted in the same way as the pulses generated by Geiger tubes. The later developments have followed from further improvements in photomultiplier sensitivity, together with the discovery of additional substances, including various liquids, which can be employed as phosphors.

The discharge mechanism in proportional and Geiger counters depends fundamentally on the separation of electrons and positive ions formed from

the molecules of a gas by interaction with ionizing radiation. The interaction that occurs in a phosphor is rather different, since electrons are not necessarily displaced altogether from atoms in the path of the radiation, but instead the energy level of some of the atoms is temporarily raised. When such excited atoms return to their ground state, the energy imparted to them in the interaction is emitted as photons of electromagnetic radiation. In fluorescent or phosphorescent substances the wavelength of this radiation is in the ultraviolet or visible part of the spectrum. Any material, liquid or solid, which can readily form excited atoms is a potential phosphor, but obviously only those materials that are transparent to the photons produced within them are of practical value for scintillation counting. Another criterion that needs to be satisfied for a substance to be a good phosphor is that the period of light emission should be very brief. In general, this is easily achieved, and the resolution times of scintillation counters are appreciably less than those of gas chamber counters, being well below 1 μ sec for most inorganic phosphors, and shorter still for organic phosphors. In order to detect γ -rays with maximum efficiency it is necessary for the radiation to traverse a substantial thickness (of the order of several cm) of the phosphor, so that each γ -photon has a good chance of giving rise to at least one flash of light somewhere within the phosphor. A further requirement which somewhat limits the choice of inorganic phosphors is therefore that it should be possible to obtain them as large clear crystals. The most commonly used inorganic phosphors are calcium tungstate, zinc sulphide and sodium or potassium iodide. The first of these substances is an efficient phosphor in its naturally occurring pure state (scheelite), but the phosphorescence yields of the others can be considerably increased by the presence of certain impurities known as 'activators'; thus zinc sulphide crystals can be activated by traces of manganese, while sodium iodide is usually thallium-activated. The function of the activators is to produce imperfections in the crystal lattice which help to ensure that the excited atoms return to the ground state by emitting radiation in the visible spectrum, rather than by handing the energy on to neighbouring atoms in such a way that in the end it merely appears as heat. Various organic compounds are also good phosphors, notably (in solid form) naphthalene or anthracene. For liquid scintillation counting of low-energy β -emitters, solutions of terphenyl in xylene or of diphenyloxazole in toluene have given the best results.

The components of a complete scintillation counting system are as follows:

(1) Container for the isotope sample. If the sample is in liquid form, as it often is when working with γ -active isotopes, it may either be placed in a small cylindrical vessel fitting into an annular slab of phosphor or placed in an annular vessel which surrounds a cylindrical slab of phosphor.

(2) A thick, light-tight lead castle enclosing sample, phosphor, photomultiplier and pre-amplifier. For very low-level β - or γ -counting, traces of radioactive impurities in the lead may give unacceptably high background counts, which can be reduced by using instead a massive iron shield and an annular tank of mercury surrounding the phosphor (see Johnston¹⁵).

(3) An efficient light connection between the phosphor and the window of the photomultiplier in order to reduce internal reflections at the surface of the phosphor to a minimum. For special applications where the phosphor

has to be at a distance from the photomultiplier (e.g. where the phosphor is used as a probe which can be inserted into living tissues containing γ -active isotopes) light can be piped from the phosphor along an internally polished metal tube, or along a light-guide consisting of a rod of transparent plastic material.

(4) A photomultiplier tube (see Chapter 28) whose spectral sensitivity matches the wavelength of the light produced in the phosphor.

(5) A well stabilized high-voltage supply (see Chapter 37) and a chain of high-stability resistors to provide suitable voltages for the electrodes of the photomultiplier tube. The overall gain of the photomultiplier varies as a high power of the applied voltage, so that good stabilization of the supply is essential.

(6) A linear pulse amplifier to make the output pulses from the photomultiplier large enough to trigger a scaling circuit, or to be fed to a discriminator circuit.

(7) A scaling circuit (see Chapter 41) with adjustable threshold.

(8) For some applications it is advantageous to use a single-channel analyser (a 'kick-sorter') between the amplifier and scaler. This consists of a double discriminator circuit, arranged so that only pulses whose height falls between the two discriminator settings are allowed to reach the scaler, smaller and larger pulses being rejected. Since, as in a proportional counter, pulse height depends on the energy of the radiation emitted by the isotope, such a device may permit discrimination against background radiation or between radiation emitted by two different isotopes. An appropriate circuit is described by Farley¹⁶.

(9) Another refinement for lowering the background in liquid scintillator low-level β -counting is to use two photomultipliers mounted on opposite sides of the sample container, and after feeding their outputs to separate amplifiers and single-channel analysers to use a coincidence circuit to count only those pulses which appear simultaneously in both channels (see Arnold¹⁷).

The pulse height given by a scintillation counter varies with: (a) the voltage applied to the photomultiplier tube, and (b) the energy of the exciting radiation. If a scaler with variable threshold is employed so that all pulses below a certain size can be rejected, the counting rate for a given isotope sample is affected both by the threshold setting and by the operating voltage. The curve relating counting rate to voltage does not show a plateau like the characteristic curve of a Geiger tube (*Figure 31.7*), but since, in general, the background pulses differ in size from those due to the sample, a setting of the controls can usually be found which represents the optimum compromise between high sensitivity and high background. By using a single-channel analyser it is possible to discriminate still better against background pulses. Good discrimination is fairly easy to achieve for isotopes giving strong β - or γ -radiation, but for weak radiation the gain of the photomultiplier tube may have to be increased to the point where pulses originating from the emission of thermal electrons at the photocathode begin to be recorded. Unless special measures are taken this sets a limit to the energy of the radiation which can usefully be detected. However, by such stratagems as cooling the photomultiplier to -20°C and

using coincidence circuitry, even the 0.018 MeV β -particles given by tritium have been recorded successfully with liquid scintillators.

CHOICE OF EQUIPMENT FOR TRACER EXPERIMENTS

It should by now be clear that the equipment needed for a tracer experiment is determined partly by the nature of the experiment and partly by the characteristics of the isotope concerned. When working with an isotope giving strong β -radiation a wide choice of measuring techniques is available. Geiger or proportional counter tubes will probably be the most convenient detecting devices, and samples can be counted in either solid or liquid form with roughly the same sensitivity. It is also possible to devise methods for measuring the amount of radioactivity in living tissues, both isolated and *in vivo*. For isotopes giving weaker β -radiation, the sensitivity of liquid counters and of measurements *in vivo* is apt to become excessively small, but if a γ -ray is also emitted, as for example in the case of ^{131}I , scintillation counting of the γ -radiation is likely to be the most efficient technique to use. The choice of counting methods only becomes really seriously circumscribed for those isotopes (e.g. ^3H , ^{14}C , ^{35}S) which emit weak β -radiation without any accompanying γ -radiation, and in experiments where maximum sensitivity is essential. Johnston¹⁵ has given a useful account of the relative merits of the various gas chamber and liquid scintillator techniques for very low-level β -counting.

REFERENCES

- ¹ KAMEN, M. D. *Isotopic tracers in biology*, 3rd ed. New York; Academic Press. 1957
- ² SACKS, J. *Isotopic tracers in biochemistry and physiology*, 1st ed. New York; McGraw-Hill. 1953
- ³ SACKS, J. Tracer techniques: stable and radioactive isotopes. In *Physical techniques in biological research*, Vol. 2, ed. by G. Oster and A. W. Pollister. New York; Academic Press. 1956
- ⁴ SIRI, W. E. *Isotopic tracers and nuclear radiations*, 1st ed. New York; McGraw-Hill. 1949
- ⁵ WHITEHOUSE, W. J. and PUTNAM, J. L. *Radioactive isotopes*. Oxford University Press. 1953
- ⁶ TAYLOR, D. *The measurement of radio isotopes*. London; Methuen. 1951
- ⁷ SHARPE, J. *Nuclear radiation detectors*. London; Methuen. 1955
- ⁸ DONIACH, I., HOWARD, ALMA and PELC, S. R. *Progress in Biophysics* Vol. 3. London; Pergamon Press. 1953
- ⁹ BOYD, G. A. *Autoradiography in biology and medicine*. New York; Academic Press. 1955
- ¹⁰ ROSSI, B. B. and STAUB, H. H. *Ionization chambers and counters*, 1st ed. New York; McGraw-Hill. 1949
- ¹¹ WILKINSON, D. H. *Ionization chambers and counters*. Cambridge University Press. 1950
- ¹² COOK, R. H. and KEYNES, R. D. *J. sci. Instrum.* 30 (1953) 141
- ¹³ KEYNES, R. D. *J. Physiol.* 114 (1951) 119
- ¹⁴ VEALL, N. *Brit. J. Radiol.* 21 (1948) 347
- ¹⁵ JOHNSTON, W. H. *Science* 124 (1956) 801
- ¹⁶ FARLEY, F. J. M. *J. sci. Instrum.* 31 (1954) 241
ARNOLD, J. R. *Science* 119 (1954) 155

APPENDIX

Explanation of operation of Figure 31.5

Before discussing the detailed operation of this circuit (which is shown in simplified form) it is perhaps as well to consider why something more elementary would not suffice. The requirement of the device is that upon the receipt at the input terminal of a brief negative-going pulse of above a certain size, a large negative-going square wave of controllable duration shall be generated at that terminal.

Clearly some kind of flip-flop is required. Suppose we set up the circuit—already discussed in Part I—of *Figure 31.10*. V_1 is normally cut off, V_2 conducts heavily and we attempt to trigger the arrangement by a negative pulse applied to V_1 anode: then we bump up at once into at least three difficulties. The first is that there is no simple way of varying the threshold

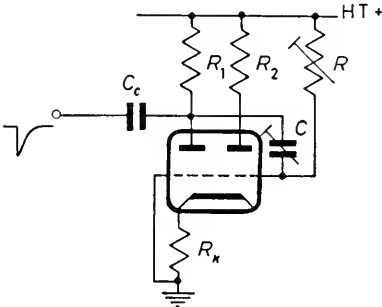


Figure 31.10

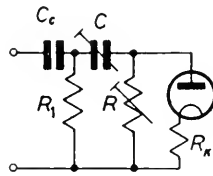


Figure 31.11

level of the device which does not also interfere with the duration and amplitude of the square wave generated. The second is more serious. The input impedance seen by the trigger pulse has the equivalent circuit of *Figure 31.11*. The shunt elements R_1 and R may be of quite high resistance, but the resistance of the conducting diode (representing the grid and cathode of V_2) in series with R_k is of the order of a few thousand ohms only. The effective internal resistance of the Geiger tube, regarded as a generator, would be much higher. Consequently only a small proportion of the open-circuit Geiger tube output would be available to trigger the circuit, probably insufficient for reliable operation.

Suppose we return R , not to $HT+$ but to a potential only just sufficiently above earth to hold V_1 properly cut off. It may then be possible to arrange that V_2 grid is sufficiently negative with respect to the cathode for inappreciable grid current to flow, in which case the input resistance of V_2 is high and the circuit will trigger satisfactorily. Unfortunately we run into a new difficulty. V_2 grid wave form is sketched in *Figure 31.12* for the two cases ' R returned to $HT+$ ' and (the time constant being appropriately modified) ' R returned to a less positive potential'. It is clear that variations in the

ASSAY OF RADIOACTIVITY

point at which conduction begins in V_2 due to valve aging produce a much more serious effect in the latter case. We see then that our flip-flop requires that R be returned to a moderately positive potential when the circuit is at rest, in the interests of high input resistance, and to a higher potential during the flip period, in the interests of precision of timing.

Our simple circuit is open to a third objection—that it is not perfectly quiescent and ready for a further flip immediately after it has flopped.

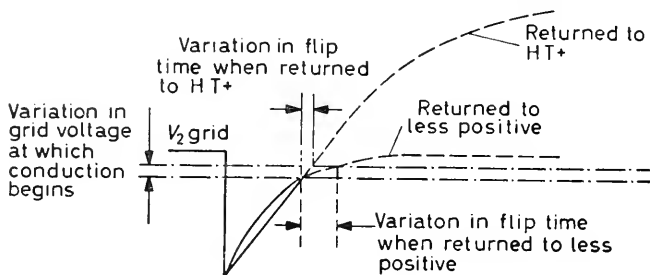


Figure 31.12

This is because of the time taken for the charge on C to reach its resting value via grid current in V_2 (t in Figure 31.13). t may be reduced by providing an additional charging path via a diode.

We are now in a position to discuss the operation of the circuit actually used (Figure 31.5). The flip-flop (V_2 and V_4) is not of the cathode-coupled variety. The positive feedback path is from anode of V_2 to grid of V_4 and from anode of V_4 via a cathode follower to the screen of V_2 . When the

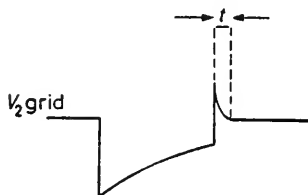


Figure 31.13

circuit is quiescent V_4 conducts and V_2 is cut off. C , R_2 and R_3 form a compensated direct coupling and R_2 and R_3 are chosen such that cathode-follower grid and cathode are approximately at earth potential. Since R_1 , corresponding to R in Figure 31.10, is returned to this line, V_4 conducts but not too hard and the input resistance is high. Since the screen of V_2 is connected to this line and the grid of V_2 is returned via V_{3A} to a yet more negative point at P , V_2 is cut off.

The magnitude of input pulse required to flip the circuit is determined by the setting of P . On the arrival of a pulse of sufficient size, V_4 is cut off and V_2 is turned on, via the cathode-follower, both by a 100 V positive-going change applied to its screen and by the removal of the holding-off bias hitherto applied to the control grid via V_{3A} . The circuit remains flipped for a time determined by the time-constant R_1C , and because R_1 is

APPENDIX

now returned to a point at 100 V positive the timing accuracy is good. The circuit then flops, rapidly returning to the quiescent state because R_1 is shunted by the diode V_{1B} . In this manner a 240 V negative-going square wave is generated at the anode of V_2 and is fed back via C_2 to the Geiger tube.

The load of V_2 is tapped so that only a fraction of the 240 V is passed on to the scaling equipment. This prevents loading of the circuit by e.g. the output cable capacitance. The reduced square wave appearing at the junction of the 3.9k and 56k resistors is differentiated and either the negative- or positive-going component is suppressed by the diode V_{1A} , leaving a monophasic output of the required polarity to operate the scaler.

ED.

VISUAL INDICATORS

P. E. K. DONALDSON

Visual indicators are those devices which permit the magnitude of electrical quantities to be read off, usually by converting them into mechanical movement along a scale. The most important are: the meter, for measuring quantities which change only slowly, or not at all; and the cathode ray tube, which displays quantities which change rapidly, usually in a system of rectangular co-ordinates in which time is the independent variable.

METERS

In this section we are not concerned with galvanometers, in which the over-riding consideration is sensitivity, or with precision meters, in which accuracy is the principal object. We are interested in the familiar instrument, of the grade known as 'Industrial', which is mounted on electronic equipment as a fixture.

The relevant meter movements are the electrostatic, the moving-iron and the moving-coil, of which the latter is the most important. Hot wire, dynamometer (wattmeter), ratiometer (Megger) and frequency meters are not of great interest to the electrophysiologist; the latter is for checking the frequency of the supply mains, and would not be appropriate for measuring, for example, the spike-discharge rate of a nerve.

Electrostatic voltmeter

The structure of an electrostatic movement resembles an air-dielectric variable capacitor. The potential difference to be measured is applied between the two sets of plates, which are in consequence attracted together against the action of a spring, to produce a movement dependent upon the potential difference. The application of the voltage to the instrument displaces a charge round the circuit so that one set of plates is charged to $+q$, and the other to $-q$. Since the attractive force is proportional to minus the product of the charges on the two sets of plates, $= q^2$, the force is proportional to the square of the applied voltage and the instrument is 'square-law'. Two important properties follow:

(1) The scale of the instrument will be cramped near zero and extended near maximum, though the effect can be to some extent offset by special shaping of the plates.

(2) If the applied voltage be alternating, or direct with an added ripple, the instrument reads true R.M.S. values.

Electrostatic voltmeters of the industrial grade are not very sensitive devices and commonly have full scale readings of the order of kV. This,

METERS

coupled with the fact that they do not consume a steady current, makes them appropriate for monitoring the values of EHT power supplies, where the regulation is commonly so poor that the current demands of a meter of some other type would produce a reading much too low.

Moving-iron meters

In one form of these, two pieces of soft iron are magnetized so that they repel one another, under the influence of a solenoid which encircles them both (*Figure 32.1*). The intensity of magnetization produced in each is proportional to the current in the solenoid, and the repulsive force is proportional to the product of the two intensities. One piece of iron is fixed,

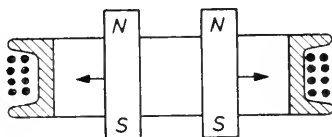


Figure 32.1

and the other is pushed away from it, against the action of a spring, to deflect the pointer. Since the repulsive force is proportional to the square of the current in the coil, the moving-iron movement is another square-law device, may be used on alternating currents and reads true R.M.S. values. The scale is cramped on the left and extended on the right, though matters may be improved by suitable kinematic design.

Moving-iron movements are not very sensitive devices, and the field of application is therefore in the measurement of appreciable currents, say greater than 50 mA. A certain firm produce a moving-iron movement with full-scale deflection 5 mA, but to achieve this a large number of turns of wire is necessary on the solenoid. The wire has therefore to be thin, and the resistance is 12,000 ohms. Clearly the introduction of such a resistance into many circuits in order to measure the current flowing will seriously upset the status quo. Similarly, when voltmeters are made by connecting the movement in series with a suitable resistor, the current drawn may pull down the voltage being measured very seriously unless the source is of low internal resistance. Hence, moving-iron voltmeters are suitable for checking the supply mains, or the outputs of a mains transformer, but would be quite unsuitable for finding the anode potential of a valve in an R-C coupled amplifier.

The value of a moving-iron movement in reading true R.M.S. may be illustrated by an example: suppose a 6 V motor car headlamp is being used as a light source and is fed from the mains via a constant-voltage transformer. The colour temperature of the lamp depends upon the R.M.S. applied voltage. The output of the constant-voltage transformer is not sinusoidal in form. Then a moving-iron voltmeter across the lamp terminals is the only appropriate meter movement; for an electrostatic instrument would be two orders too insensitive, and a moving-coil and rectifier instrument—which reads the average of the modulus of the input—would give

a false indication. Whilst the error as a percentage in volts might not be very great, since light output changes very rapidly with applied voltage, the error in light output could be serious.

Moving-coil meter

In the moving-coil meter the field due to the moving-coil reacts with a steady field due to a powerful permanent magnet to twist the coil and deflect the pointer against the action of a spring (*Figure 33.25*). The torque depends on the product of the coil field and the permanent magnet field, and by making the latter large a very sensitive movement can result. The pointer deflection is proportional to the average coil current, and the scale is linear.

In a range of moving-coil ammeters and voltmeters of a particular size and made by a particular manufacturer it is possible that they all contain the same basic movement, perhaps 1 mA full-scale deflection and of 100 ohms coil resistance. Higher current ranges are then obtained by the addition of appropriate shunts, and the voltmeter range by suitable series resistors. The sensitivity of voltmeters is often given in terms of 'ohms per volt', a figure which is merely the reciprocal of the current for full-scale deflection. Thus the 1 mA full-scale deflection movement yields a range of voltmeters of 1,000 ohms/V, and the resistance of the 300 V full-scale deflection member of the range would be 300,000 ohms. Since the resistance of a voltmeter should be as high as possible, the ohms/V figure gives a measure of the virtue of the instrument in this respect.

For lower full-scale deflection currents than 1 mA it is necessary to have a better movement, i.e. one with a stronger magnet, more turns of thinner wire on the moving-coil, more friction-free pivots and weaker control springs. The device is therefore more fragile both mechanically and electrically. A 1 mA movement is probably as mechanically robust as many other electronic components, but a 50 μ A F.S.D. instrument should be handled with circumspection. The most sensitive industrial grade movement known to the writer is the Taylor, model 500, 0.5 μ A. He has no experience of one of these, but would suggest it be afforded the same sort of care in handling one would devote to, say, a large cathode ray tube.

Sizes of meters

Meters are described by the arc length of the scale, in inches. The 2, 2½ and 3½ in. sizes are commonest, but tiny 1 in. models are available where panel space is limited, and 5 in. movements may be used where the scale has to be read with great precision.

Accuracies of meters

Accuracies for meters are laid down in a rather complicated British Standard Specification, BS 89/54. The limits of error are expressed as percentages of the full-scale deflection at 20°C. In the important case of moving-coil meters not of the multi-range type, the following errors are allowed over pointer indications in the range 10 to 100 per cent of full scale.

METERS

<i>Scale length</i>	<i>Limits of error %</i>
1½ in. up to (but not including) 2 in.	2.5
2 in. up to (but not including) 3¼ in.	2
3¼ in. and over	1

Similar accuracies are decreed for moving-iron voltmeters, and moving-iron ammeters are allowed slightly larger errors, but in both cases the limits only have to be met between 20 and 100 per cent of full scale.

Non-linear moving-coil movements

Particularly in war surplus equipment, one occasionally comes across moving-coil movements having non-linear relationships between current and deflection, often approximately logarithmic, and these are often useful.

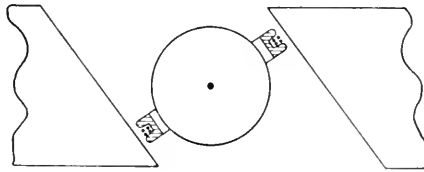


Figure 32.2

In the case of end-zero meters, the law may be obtained by shaping the permanent magnetic pole pieces as shown in Figure 32.2. With the coil in the position shown the permanent-magnet field encountered is strong, but weakens as the coil turns clockwise. This gives a scale which is expanded on the left and cramped at the right. Centre-zero movements for null-detection may be made sensitive in the region of the null-point by using pole pieces as shown in Figure 32.3, or by using conventional pole pieces and an elliptical central armature (Figure 32.4).

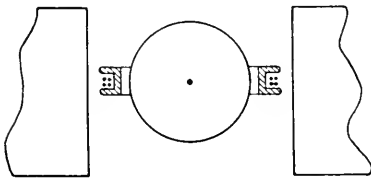


Figure 32.3

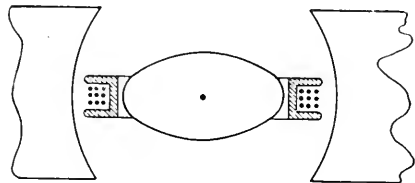


Figure 32.4

Forms of construction such as these afford mechanical protection to the movement from excessive currents, since a current much in excess of normal is required to fling the pointer hard enough on the end stop to cause damage: they cannot, of course, prevent the coil itself from accidental burn-out.

VISUAL INDICATORS

Electrical protection of moving-coil movements

Meters may be protected from overload by the connection in series of Wollaston wire fuses, and the construction of these has been described by Strafford¹. Alternatively, use may be made of the non-linear resistance properties of semi-conductor diodes (*Figure 32.5*). When a potential difference across the movement is within the normal range, neither diode is switched completely to full-conduction, and the diode resistances are therefore high; most of the current supplied passes through the meter. Under overload conditions, however, one or other of the diodes conducts strongly, the resistance is low and most of the applied current flows through that diode instead of through the meter. The scale is no longer linear, so in cases where both are applicable the choice between fuse protection

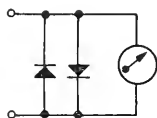


Figure 32.5

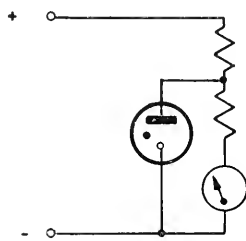


Figure 32.6

and rectifier protection is determined by whether a linear or a logarithmic scale is required. For details of rectifier protection methods Scroggie² and de Gruchy³ may be consulted. Scroggie also quotes a method for the electrical protection of meters reading voltages above 100 or so, due to T. E. Burnup. A cold cathode diode is connected across the meter and an appropriate fraction of the series resistor, such that the tube strikes just before the applied voltage becomes excessive (*Figure 32.6*).

Use of meter amplifier

The cost and fragility of moving-coil movements rises steeply as the full-scale deflection current falls. To obtain a meter with a full-scale deflection of a few microamps it may be cheaper, and a robusiter job be obtained, to consider the use of transistor amplification. Johnson⁴ has described a push-pull transistor amplifier, small enough to fit into the meter case, which converted a $50 \mu\text{A}$ F.S.D. movement into an instrument giving a full-scale deflection for less than $2 \mu\text{A}$. Apparently no difficulty was experienced from drift.

A group of meters is shown in *Plate 32.1*. These are all of the 'flush-mounting' type: (A) is 1 in., end-zero, round, moving coil; (B) is 2 in., end-zero, square, moving coil; (C) is $2\frac{1}{2}$ in., centre-zero, round, moving coil; (D) is 3 in., end-zero, round, moving iron; and (E) is $4\frac{1}{2}$ in., end-zero, rectangular, moving coil.

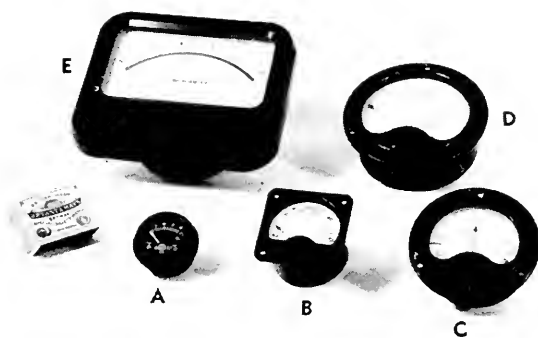


Plate 32.1 Group of 'Industrial grade' meters



Plate 32.2 Magnetic deflector coil assembly

'MAGIC EYE' NULL-DETECTOR

NEON LAMP INDICATOR

An important application for meters is the use of centre-zero movements as null-indicators in the outputs of, e.g., bridges. Unfortunately it often happens that the impedance looking back from the meter into the bridge is much higher than the resistance of the meter itself; that is, the meter is mismatched to the bridge, power transfer conditions are poor, and the system is insensitive.

The usual procedure in these cases is to interpose a pair of cathode followers between bridge and meter, as shown in *Figure 32.7*. The full open-circuit bridge output is then transferred to the meter terminals. However, if one is going to introduce a couple of valves, it may be worth considering a simple null-detector due to Collinson⁵ in which, instead of using a meter, resort is made to the ability of the eye to match brightnesses accurately. Small neon lamps (Hivac CC 11L) are included in the anode

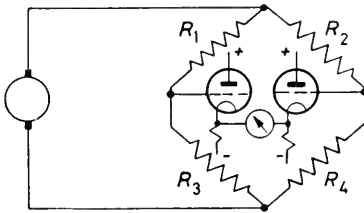


Figure 32.7

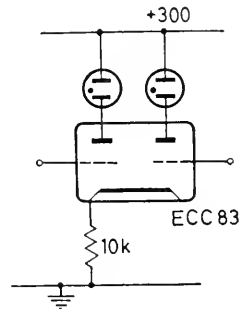


Figure 32.8

circuits of a long-tailed pair (*Figure 32.8*). When the grids are at the same potential the neons should glow at equal brightness. The input impedance of the device is, of course, very high, and it has the merits of rapid response and indestructibility. Per volt of potential difference at the input, one lamp brightens, and the other dims, by about 0.05 lm. With both grids earthed, each lamp emits about 0.06 lm. Hence with one grid made $\frac{1}{2}$ V positive to earth and one $\frac{1}{2}$ V negative, the light outputs become 0.11 and 0.01 lm, so the device is quite sensitive.

'MAGIC EYE' NULL-DETECTOR

Another null-detector⁶, employing single-sided circuitry, uses the familiar 'magic eye' valve, generally used for facilitating the tuning of radio receivers. Like the neon lamps these valves have the advantage over meters of instant response and electrical indestructibility. The circuit is shown in *Figure 32.9*. With zero input voltage the shadow angle of the magic eye is 85 degrees. An input of either polarity then narrows the shadow angle; 1 V positive reduces it to 55 degrees and 1 V negative to 30 degrees.

VISUAL INDICATORS

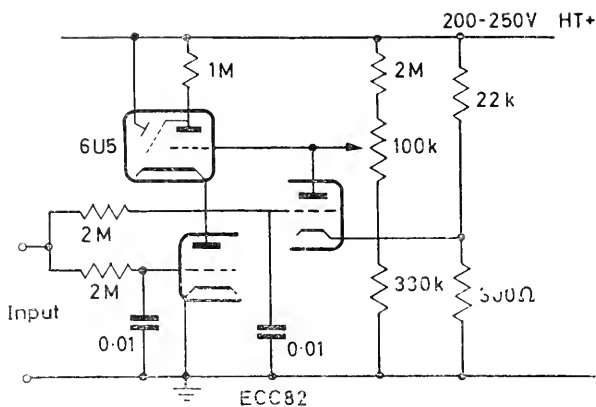


Figure 32.9

DUDELL OSCILLOGRAPH

The Duddell oscillograph is virtually a mirror galvanometer in which the suspension has been stiffened to sacrifice sensitivity for speed of response. A beam of light reflected from the mirror may be used to expose a moving photographic film, so that a record is obtained in which the movement of film produces the time base. To produce a visual display, however, the oscillograph beam has to be reflected on to a screen by a second, rotating mirror; not only is the method cumbersome, but also the upper limit of frequency response of even the fastest Duddell oscillographs is a few thousand cycles per second. For these reasons cathode ray oscillography has largely supplemented mechanical methods for visual indication.

Matthews oscillograph

Devised by B. H. C. Matthews in 1927, this is an electromagnetic moving-iron instrument which, like the Duddell, deflects a mirror and hence swings a reflected beam of light. It was designed specifically for electrophysiological recording, and was widely employed until the late 30s, when cathode ray tubes became cheap and reliable. A few are still in use. The frequency response extends to about 6,000 cycles per second.

CATHODE RAY TUBE

Cathode ray tubes are of three types, electrostatic, magnetic and hybrid.

Electrostatic cathode ray tubes

These mostly have round screens, in diameters ranging from 1 to 12 in., of which the great majority are between 4 and 6 in. A few types have screens of other shapes: one model (20th Century S6 sq.) has a square screen of 6 in. diagonal, and one or two firms make a tube having a rectangular $6 \times 1\frac{1}{2}$ in. screen. Phosphor materials have been discussed in Chapter 28.

CATHODE RAY TUBE

The construction of an elementary type of electrostatic tube is shown in *Figure 32.10*. Electrons are emitted from an indirectly heated cathode and are attracted to a relatively positive first anode, which is a disc perforated by a small central hole. The emission rate is controlled by controlling the space charge around the cathode with a relatively negative electrode which, by analogy with a triode valve, is called a grid, though in fact it is cylindrical. The amount of grid bias used determines the brightness of the trace. Some of the electrons emitted pass through the hole in

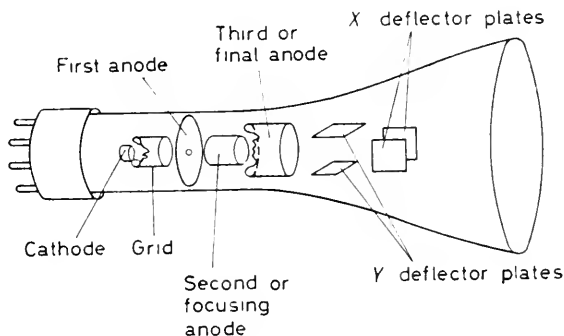


Figure 32.10

the first anode and come under the influence of cylindrical anodes 2 and 3. The 3 anodes together form an electrostatic lens, which focuses the beam so that an image of the hole in A_1 appears on the screen. Often A_1 and A_3 are connected together and to, say $+2,000$ V with respect to the cathode. A_2 is fed from a focus potentiometer with, perhaps, $+700$ V.

The focused beam then comes successively under the influence of the Y and X deflector plates. As normally used, the former impart a vertical velocity to the electrons, bending the beam up or down, and are fed with the signal to be displayed. The X plates impart lateral velocity to the electrons, bending the beam sideways, and are usually fed with the output of the time-base generator.

Finally the deflected beam hits the screen, exciting the phosphor and producing a small, round, bright spot. The electrons subsequently find their way back to A_3 via leakage paths within the glass envelope which is, of course, evacuated.

Deflection sensitivity of electrostatic cathode ray tube—*Figure 32.11* shows part of a cathode ray tube, in which A_3 is positive to the cathode by an amount V_{A_3} , and a pair of deflector plates, potential difference V_D , distant apart d , and l long, bend the beam an angle θ so that the spot moves a distance D . The length of the neck of the tube is L .

$$\text{Then} \quad \frac{D}{L} = \tan \theta = \frac{\text{lateral electron velocity}}{\text{axial electron velocity}} = \frac{v_l}{v_a}$$

The kinetic energy acquired by an electron, of charge e , in being accelerated by an electric field whose boundary potentials are different by E , is eE . In our case the electron is accelerated between cathode and third anode, the

VISUAL INDICATORS

potential difference involved is V_{A_3} and the energy acquired is $e \cdot V_{A_3}$. We have, if m is the mass of an electron

$$\frac{1}{2}mv_a^2 = e \cdot V_{A_3}$$

$$v_a^2 = \frac{e}{m} \cdot 2V_{A_3}$$

On emerging from the third anode, the electrons encounter a lateral field, due to the deflector plates, of magnitude V_D/d . The lateral force on the

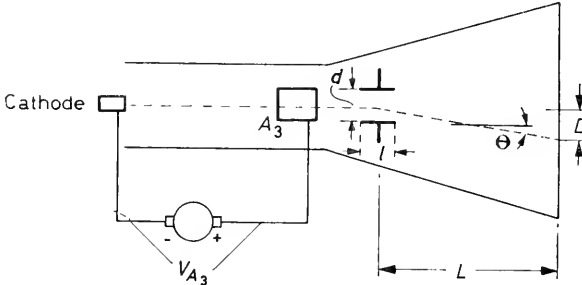


Figure 32.11

electron is $(V_D \cdot e)/d$, producing an acceleration $(V_D \cdot e)/dm$. The acceleration acts for a time t , during which the electron is between the plates; but $t = l/v_a$. The lateral velocity which the electron attains is acceleration \times time

$$= \frac{V_D \cdot e \cdot l}{d \cdot m \cdot v_a}$$

$$\therefore \tan \theta = \frac{\text{lateral velocity}}{\text{axial velocity}} = \frac{V_D \cdot e \cdot l}{d \cdot m \cdot v_a^2}$$

but

$$v_a^2 = \frac{e}{m} \cdot 2 \cdot V_{A_3}$$

$$\therefore \tan \theta = \frac{1}{2} \cdot \frac{V_D}{V_{A_3}} \cdot \frac{l}{d}$$

It follows from the above that:

(1) For a given deflection angle, a long tube produces a bigger deflection per unit deflector plate potential difference—i.e. is more sensitive—than a short one.

(2) For good sensitivity the deflector plates should be long, or close together, or both. In either case it is clear from the geometry that the maximum possible deflection angle is severely limited by the beam striking one or other deflector plate. Hence, electrostatic cathode ray tubes have to be long in relation to their screen diameter.

(3) For good sensitivity the EHT should be the lowest possible consistent with securing a spot of adequate brightness and focus. The maker's figure for cathode ray sensitivity is given in the form say, $800/V_{A_3}$ mm/V, meaning

CATHODE RAY TUBE

that if the EHT is 2.4 kV, 0.33 mm deflection is produced for a potential difference of 1 V between the deflector plates.

Satisfactory operation of these tubes is dependent upon the maintenance of the correct mean potential in the region between the pairs of deflector plates with respect to the potential of the third anode. Each pair of plates should be fed in push-pull so that the mean potential remains the same, and the mean potential should be arranged to be equal to, or not far from, the third anode potential. Cathode ray tubes are prone to certain defects, some of which are due to failure to pay sufficient attention to this point.

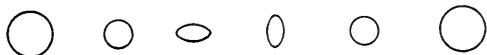


Figure 32.12

Astigmatism—This is a refusal upon the part of the spot to focus properly. On manipulating the focus control both the height and width of the spot can be made to pass through minima, but the minima do not occur at the same settings of the control; hence the spot passes through the stages shown in Figure 32.12. The trouble is caused by an unsatisfactory relationship between the mean potentials of the two pairs of deflector plates.

Trapezium distortion—This is a deflection distortion caused by failure to drive the pair of deflector plates further from the tube cathode in push-pull. If the tube has *X* plates further from the cathode, and an attempt is made to use it to draw on the screen a rectangular grid like Figure 32.13a, what in

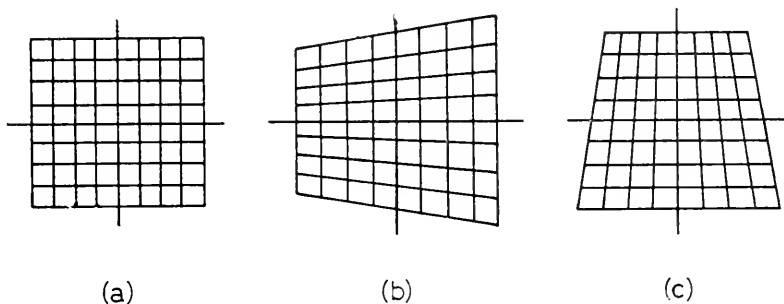


Figure 32.13

fact appears is Figure 32.13b; if the *Y* plates are further from the cathode and are asymmetrically fed, Figure 32.13c results.

Trapezium distortion does not follow the use of asymmetric drive to the deflector plates nearer the cathode, but other faults may, such as a variation in the goodness of the focus in different parts of the screen.

Briefly, the mechanism of trapezium distortion is that the unbalanced drive to the deflector plates causes the mean potential to differ from that of the final anode, setting up an additional weak electrostatic lens. The beam undergoes deflection as it passes the first pair of plates, and is then subjected to refraction *in the same plane* according to the power of the lens formed by the second pair.

Pin-cushion distortion—With electrostatic cathode ray tubes, the effect of

VISUAL INDICATORS

'pin-cushion' distortion is to produce a grid like *Figure 32.14a* or *b* dependent upon whether Y plates or X plates are nearer to the cathode: this is clear

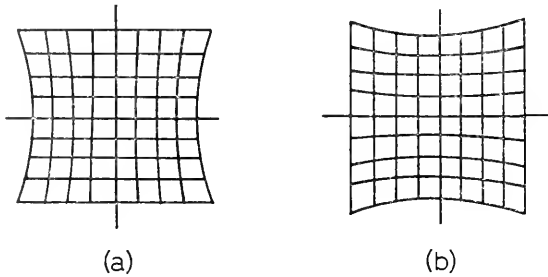


Figure 32.14

for the former case from *Figure 32.15*. When the beam leaves the region between the Y plates it encounters a fringe field possessing an axial component which retards it. In consequence it suffers more X deflection, leading to a grid of the *Figure 32.14a* type.

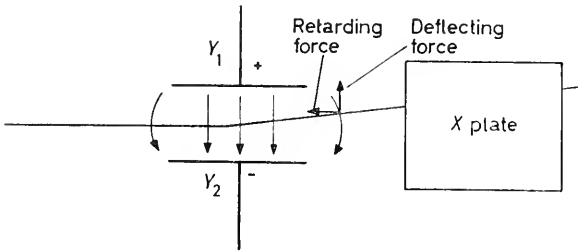


Figure 32.15

Cossor double-beam electrostatic tube—Cathode ray tube distortions may be minimized by special shaping of the electrodes and by the introduction of certain additional ones, e.g. the Fleming-Williams' curved anti-trapezium plate⁷. By the use of such techniques the Cossor double-beam tube is rendered feasible, in which the deflector plate drive is *intended* to be asymmetrical. Part of this is sketched in highly simplified form in *Figure 32.16*. The beam is divided into two, whilst still in a diffuse state, by a 'splitter'

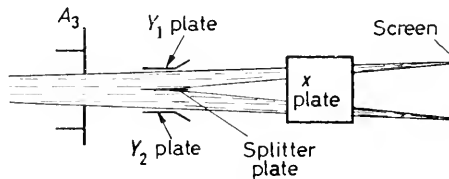


Figure 32.16

plate at A_3 potential. The upper part is then vertically deflected by asymmetric drive to the Y_1 plate, to show one quantity, and the lower is deflected by the Y_2 plate to show another. Both beams then receive lateral deflection from the X plates, to one of which the time-base waveform is connected.

CATHODE RAY TUBE

Considering that the time-base output is also asymmetric, these tubes represent, in their freedom from distortion, a remarkable achievement.

Tube supply networks—A typical ancillary circuit for feeding a conventional electrostatic cathode ray tube is shown in *Figure 32.17*. The deflection voltages are capacitor-coupled to the plates, so the time constant of the signal path circuit, CR , must be adequate to transmit the lowest waveform frequency it is required to examine. Since the X plates are also capacitor coupled to the time-base generator, it is unlikely that sweep times longer than about $\frac{1}{4}$ sec would be feasible, since an extremely large value of coupling time constant $C'R'$ would be necessary to transmit the triangular waveform

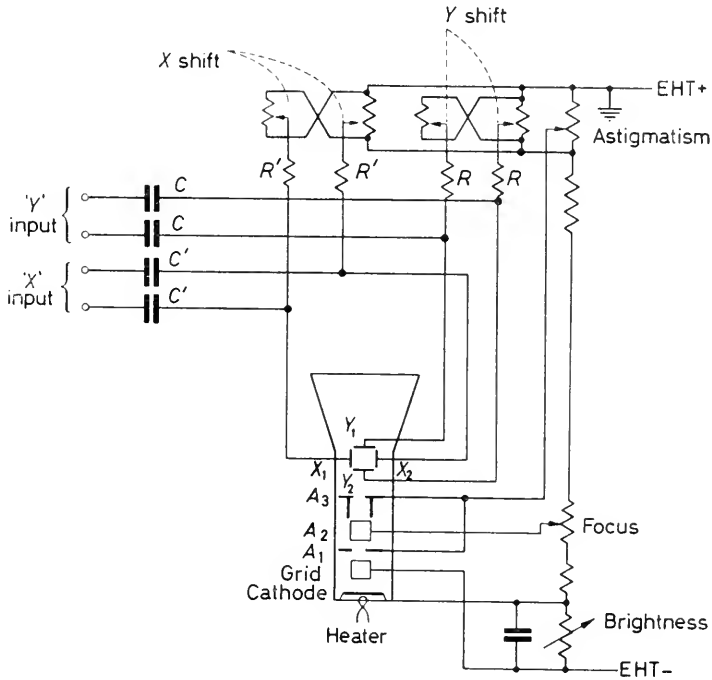


Figure 32.17

without serious distortion. Such a circuit is not therefore of much use in electrophysiological work, but it illustrates the kind of network to be found in certain commercial oscilloscopes. Notice how the shift potentials are applied in push-pull; the sliders move up or down each ganged potentiometer together. Notice also the astigmatism potentiometer for finding the optimum focusing conditions. The cathode ray tube heater is fed from a special winding on the mains transformer, insulated to withstand the EHT voltage. The capacitor across the 'brightness' variable resistor is important; the degree of smoothing provided on most EHT power supplies is sufficient to steady the inter-anode potentials, but the spot brightness is extremely sensitive to variations in grid-cathode potential, and additional smoothing is invariably required here. About $0.5 \mu\text{F}$ is generally sufficient.

The calculation of the values required in the resistor chain is quite easy.

VISUAL INDICATORS

The procedure is to choose a 'bleed' current through the chain, large compared with any of the tube electrode currents, say, 1 mA. The chain may then be treated as a simple unloaded potential divider, and the component values chosen to provide the necessary ranges of potential. In oscilloscope practice it is the rule to earth the positive EHT terminal, otherwise the coupling capacitors to the plates have to be rated at approximately the full EHT.

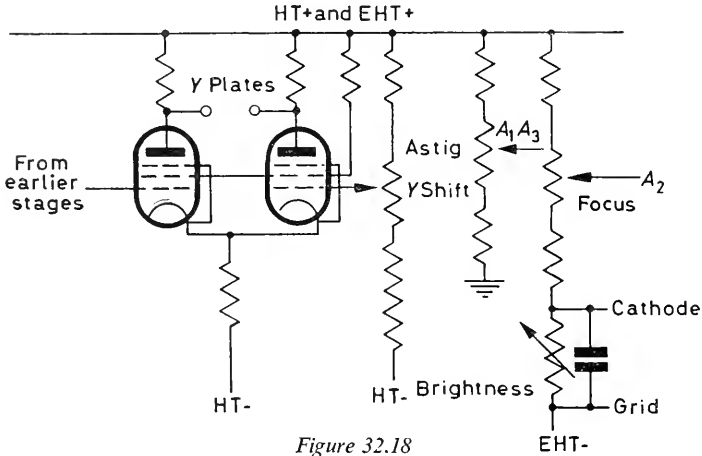


Figure 32.18

For general electrobiological work, direct coupling to the deflector plates is more satisfactory, and in this case the deflection amplifiers and the tube supply network have to be designed together. *Figure 32.18* shows a possible scheme in which a long-tailed pair forms the output stage of the Y deflection amplifier. The output of the time-base generator may be fed to the X plates

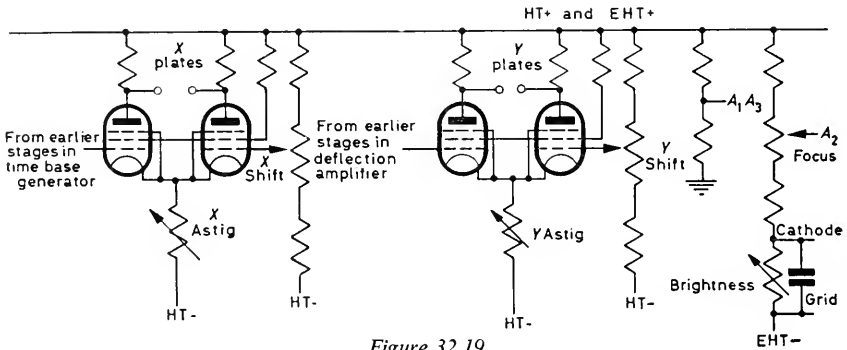


Figure 32.19

by a similar long-tailed pair, and if this is done a highly refined form of astigmatism control becomes possible in which the first and third anodes are fed from a fixed tap across the HT supply, and the long-tailed pairs are each provided with tails of variable length (*Figure 32.19*). In this way the mean potential of both pairs of deflector plates may be adjusted with respect to the third anode to achieve optimum focusability.

CATHODE RAY TUBE

However, if the time-base generator is a Miller valve, the amplitude of the output is already appreciable and no extra amplification is required beyond the provision of a second output, in opposite phase to the Miller output and equal to it, so that the X plates are fed in push-pull. To do this we use a 'paraphase' amplifier (Figure 32.20). If R_1 and R_2 are equal, we have negative feedback round the paraphase valve in which $B = 1$, so the gain is $A/(1 + A)$, where A is the gain of the latter without feedback. $A/(1 + A)$ is, of course, just less than unity, and the paraphase amplifier gain may be

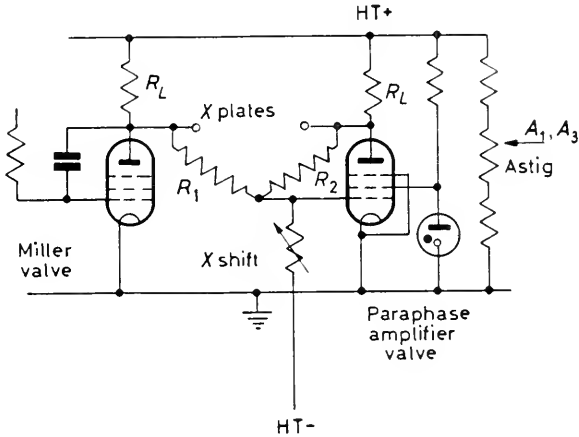


Figure 32.20

made exactly equal to unity by making R_2 rather greater than R_1 . R_1 and R_2 may reasonably be about 5 times the values of the anode loads of the valves, and the value of the X -shift control is, of course, determined by the HT negative supply available and the bias required by the paraphase valve. Screen stabilization of the latter is only necessary in the highest class of apparatus and it is often possible to omit the stabilizer tube, or even to replace the pentode by a triode, here. The performance of the circuit at high sweep speeds is improved by shunting R_1 and R_2 by equal small capacitors, say of 25 pF, to compensate for the input capacitance of the amplifier valve.

Post deflection acceleration (P.D.A.)—This is a recent development with electrostatic cathode ray tubes by which high deflection sensitivity is combined with good brightness and focus. The scheme is to use rather low A_1 , A_2 and A_3 potentials so that the electrons pass slowly between the deflector plates, and then to accelerate them in a further electrostatic field. P.D.A. electrodes may be formed by one or more annuli of deposited graphite on the inside of the flare of the tube envelope (Figure 32.21a). These are connected to potentials considerably positive with respect to A_3 and the deflector plates. For example, the G.E.C. type 1601 BCCA(F) cathode ray tube has $V_{A_4} = 3.5$ kV, $V_{A_5} = 7.5$ kV and $V_{A_6} = 10$ kV. Another, extremely elegant, P.D.A. scheme is to use a single electrode in the form of a deposited graphite spiral (Figure 32.21b) along which there is a potential gradient, so that the accelerating field is continuous rather than stepped.

VISUAL INDICATORS

The value of P.D.A. tubes lies mainly in securing an adequately bright trace when using a triggered time base and very high sweep speeds. As normally used, sweep speeds in electrophysiological work seldom exceed 6 in. in 100 μ sec, and for this conventional tubes are generally adequate. However, in using special techniques, such as multi-channel oscillography employing the voltage-coincidence method—to be described later—higher writing speeds are encountered and P.D.A. tubes are of great help.

Magnetic shielding of electrostatic tubes—Electrostatic tubes generally need to be shielded from the effects of stray magnetic fields from devices

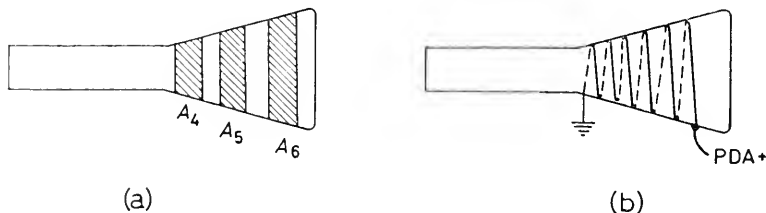


Figure 32.21

such as mains transformers. To this end, mumetal screens, which fit closely round the outside of the glass envelope, may usually be obtained from the tube manufacturers. The high permeability of mumetal is destroyed by cold working, so these screens should be subjected to as little mechanical distortion as possible when being fitted. In exceptionally bad cases one screen may prove to be insufficient; when this happens it is better to use two thin screens, concentrically arranged, than one thick one⁸.

Magnetic cathode ray tubes

Magnetic cathode ray tubes have not hitherto been much used in scientific oscillography, probably because the ancillary circuitry is more difficult to design. However, in-so-far as they are more satisfactory for television and radar, a great deal of research and development has been devoted to them and their associated components in the last 20 years, and it seems to the author that the use of magnetic tubes may now be at least considered for laboratory applications. Their advantages are:

- (1) A brighter and better focused spot.
- (2) The tubes can be much shorter for a given screen diameter. In the construction of large demonstration oscilloscopes for lecture-theatre use the equipment is not excessively bulky.
- (3) Because television serves a mass market, tubes with relatively enormous screens are readily available, whereas a comparable electrostatic tube would have to be specially made, at much greater expense. At present magnetic tubes range from models having round screens of 5 in. diameter, to rectangular screens of 21 in. diagonal. It must be admitted that this argument weighs only in considering demonstration oscilloscopes.
- (4) Because of the mass market for television, suitable pieces of ancillary equipment such as focus magnets and deflector coil assemblies are cheaply and readily available.

CATHODE RAY TUBE

(5) The advent of the transistor. The transistor is essentially a current-controlled and -controlling device, and it seems reasonable to suppose that future transistorized apparatus will employ magnetic cathode ray tubes.

Construction—The triode magnetic cathode ray tube contains merely a cathode, with associated heater, a grid to control the trace brightness, and an anode. The assembly resembles cathode, grid and A_1 in an electrostatic tube, except that in the magnetic case the anode comprises also a deposited graphite coating inside the flare of the envelope (*Figure 32.22*). The *negative*

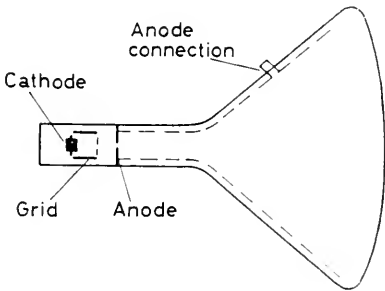


Figure 32.22

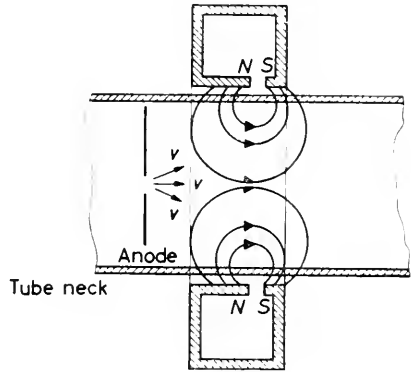


Figure 32.23

side of the EHT supply is earthed. In the U.S.A. the tetrode tube has always been preferred, and these are now common in this country also. The tetrode tube possesses two anodes: A_1 is a small metal anode near grid and cathode, and is supplied with about 300 V positive; A_2 is the graphite coating and is maintained at between 5 and 15 kV positive.

Focusing with magnetic cathode ray tubes is achieved by an axial magnetic field derived from an annular magnet (*Figure 32.23*) which fits over the neck of the tube. It can be shown that any electron emerging from the hole in A_1 with speed v will be brought back to the axis of the tube in a distance $2mv/He$ in c.g.s. units, where H is the axial focusing field strength, so that a focused spot is obtained. The experimenter does not have to concern himself with this equation, since permanent-magnet focusing assemblies may be bought, in which H is adjustable over the relevant range by mechanically varying the air gap. The only point to watch is that magnetic cathode ray tube necks are 23 mm, 35 mm or 38 mm in diameter, and it is important to buy a focusing assembly of the correct size.

The beam may be deflected vertically by passing it through a lateral magnetic field (*Figure 32.24*) or laterally by passing it through a vertical field. An exact expression for the deflection sensitivity is difficult to give and in any case does not concern us. It is sufficient to know that the deflection is proportional to the field strength and inversely proportional to the square root of the EHT. In practice one persuades a television service man to give one a burnt-out ferrite-cored deflection-coil assembly, strips out the old windings, inserts temporary new ones of known number of turns, and measures the deflection sensitivity in terms of inches of deflection produced

VISUAL INDICATORS

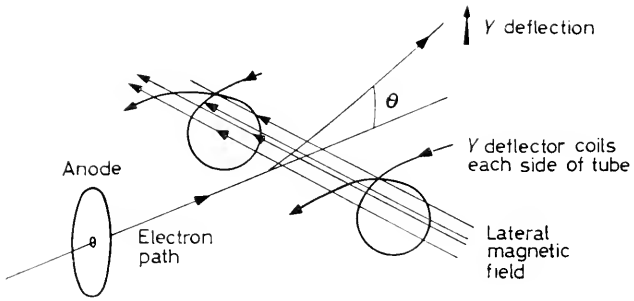


Figure 32.24

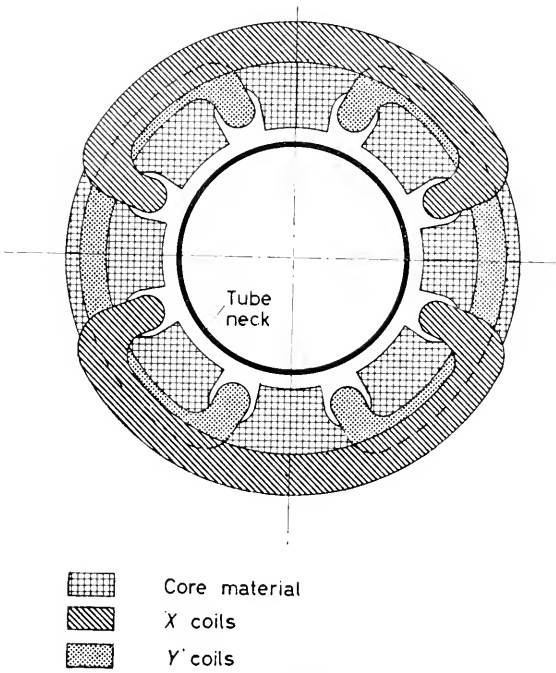


Figure 32.25

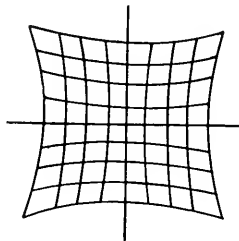


Figure 32.26

CATHODE RAY TUBE

per ampere-turn in the winding. A deflector coil assembly is typically a slotted annulus of magnetic material into which the windings are recessed, which fits over the tube neck as shown in *Figure 32.25*. The arrangement of windings may be easier to see in *Plate 32.2*, but here the annulus is not slotted. For other and simpler forms, consult, e.g. Cocking⁹. Alternatively, new slotted annuli in Ferrocube may be obtained from Messrs Mullard Ltd. (e.g., FX1154).

Magnetic cathode ray tubes proper are singularly free from faults; most of the display distortions are caused by imperfections in the deflector coil assembly. An important tube fault is:

Biaxial pin-cushioning (*Figure 32.26*). A given deflector coil current produces a certain beam deflection *angle* (*cf.* tangent of angle in electrostatic tubes) and hence if the deflection distance is to be proportional to the

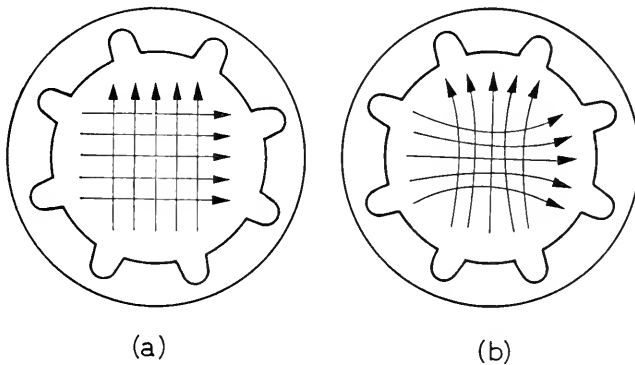


Figure 32.27

deflecting current the screen ought to lie wholly on the surface of an imaginary sphere with centre to centre of the coil assembly. Such a curved screen would be unsatisfactory, and practical screens are flatter than this, and in consequence the length of the beam increases with deflection angle, to produce greater sensitivity at the screen periphery, giving the grid of *Figure 32.26*. The effect may be allowed for by designing the deflector coils so that the fields are not uniform (*Figure 32.27a*) but are stronger near the tube axis (*Figure 32.27b*).

Deflection defocusing—This is a loss of sharpness of the spot as it is deflected away from the tube axis. It is attributable partly to the change in beam length due to the flat screen and partly to poor deflector coil design. The fringes of the deflecting field contain components which act along the axis of the tube to modify the focusing conditions (*Figure 32.28*). The remedy is to have deflector coils which are long along the tube axis, so that fringe effects are rendered as far as possible relatively insignificant; also to bend the *ends* of the deflector coils away from the tube (*Figure 32.29*).

Lack of similarity between the two *Y* deflector coils can produce a grid of the type shown in *Figure 32.30a*, and between the *X* coils, *Figure 32.30b*. Insufficient perpendicularity between the *X* and *Y* fields can produce orthogonality errors (*Figure 32.31*) though this is unlikely to occur if the coils are accurately located by fitting tightly the slots in the moulded core.

VISUAL INDICATORS

It will be clear from the foregoing that the successful use of magnetic tube displays lies largely in the careful design and construction of the deflector coil assembly. Reasonable care in winding, shaping and fitting the coils is all that is needed to minimize all faults except biaxial pin-cushioning; the necessary graded field to minimize this requires fairly advanced design, and Cocking may be consulted, though in the writer's opinion the effect is not a very serious one, and may often be ignored.

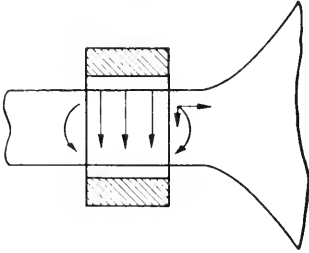


Figure 32.28

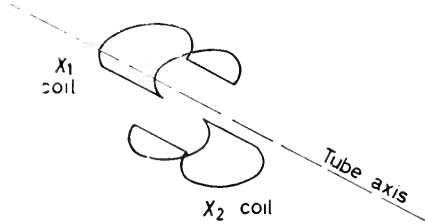


Figure 32.29

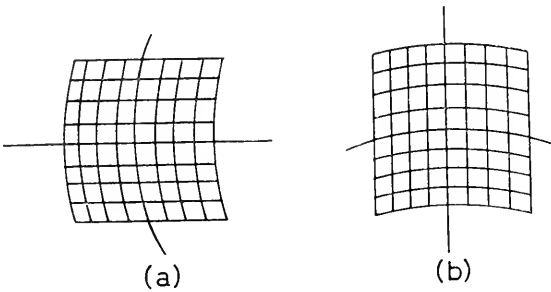


Figure 32.30

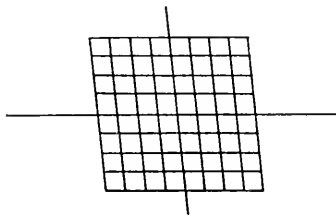


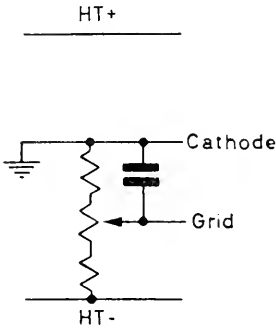
Figure 32.31

Ancillary circuitry—With magnetic cathode ray tubes it is helpful if a negative HT supply is available, since the cathode can then be earthed, the heater can be excited from the same a.c. supply as feeds the other valves, and no strain is imposed on the heater-cathode insulation. *Figure 32.32* is suitable for a triode tube and *Figure 32.33* for a tetrode. If there is no negative supply, *Figures 32.34* and *32.35* are possible, but the heater-cathode voltage must be watched, and a separate heater winding in the mains transformer may be required for the tube. No bleeder is necessary across the EHT

CATHODE RAY TUBE

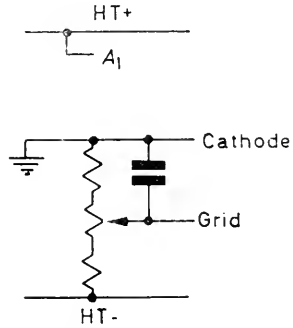
supply (except a very high value resistor to discharge the capacitors after switching off, provided in the interests of safety) which in consequence has only to be able to supply a current of the order of $100 \mu\text{A}$.

If the focusing is by permanent magnet, all one has to do is to adjust the gap as required. If electromagnetic focusing is used, the coil may—if of low resistance—be employed as an additional smoothing choke in the HT power



Triode tube

Figure 32.32

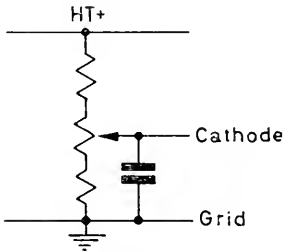


Tetrode tube

Figure 32.33

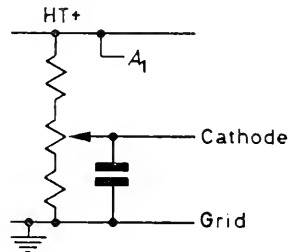
supply, and the current through it controlled by a variable shunt resistor. If it is of high resistance, it should be connected across the HT supply with a variable series resistor. Various other small magnets are used with magnetic tubes. One is the ion-trap magnet.

Ion-trap magnet. In the expression for electrostatic deflection the ratio e/m cancels out; that is, all charged particles follow the same path irrespective of their mass. In a practical tube heavy ions, both positive and negative,



Triode tube

Figure 32.34



Tetrode tube

Figure 32.35

are present. In electrostatic tubes the negative ions are also projected at the screen and strike it, at any instant, in the same place as the electrons.

The deflection expression for magnetic tubes contains e/m , in such a way that heavy particles undergo less deflection than light ones. The effect is that the heavy negative ions bombard a restricted area of screen near the tube axis, and eventually damage it and produce 'ion burn', an insensitive dark patch in the middle of the screen.

VISUAL INDICATORS

Ion burn may be combated in tetrode tubes by an elementary application of mass spectrometry. The electron 'gun'—consisting of grid, cathode and first anode—is inclined to the axis of the tube (*Figure 32.36*). Under the influence of a lateral field—from a small assembly of permanent magnet and pole pieces, which slips over the tube neck—the electrons are diverted into their proper path along the axis. The ions are not; they hit the side of the

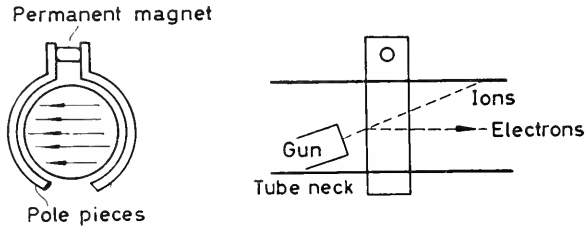


Figure 32.36

tube and are eventually collected by A_2 . Adjustment of the ion-trap magnet is a simple matter of searching for the position which gives the brightest trace.

Shift magnets. Rather similar small assemblies may be bought which provide a transverse field of variable strength and direction to provide beam 'shift' facilities. From *Figure 32.37* it is clear that the device is similar to the ion-trap, but the permanent magnet may be turned as shown between the pole pieces so that the strength of the field between them follows a cosine

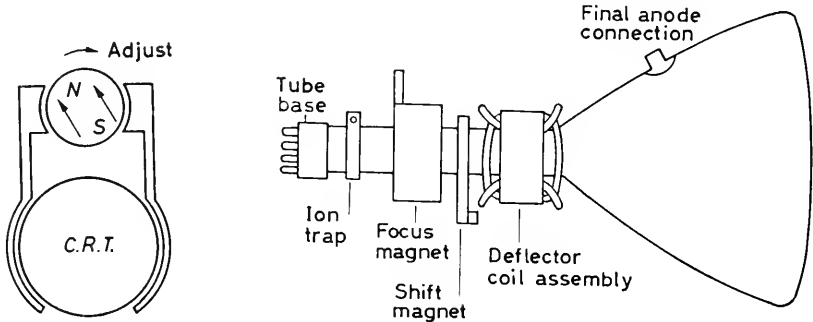


Figure 32.37

Figure 32.38

law. By rotating the device as a whole round the tube neck the direction of shift is controlled. The magnitude of the shift is determined by the position of the rotatable permanent magnet within the device itself.

The order in which the various ancillaries are assembled is shown in *Figure 32.38*.

Deflection amplifiers

The direct-coupled deflection main-amplifier of the author's demonstration oscilloscope is shown in *Figure 32.39*. A long-tailed pair is used as a phase-splitter to feed a push-pull power output stage in class AB. The frequency response is 6 dB's down at 10 kc/s. Using a 17 in. diagonal rectangular screen

CATHODE RAY TUBE

tetrode cathode ray tube, with a second anode potential of 10 kV, the deflection sensitivity is one in. per volt. Each deflector coil comprises two windings, *A* and *B*, fed from the two sides of the output stage. Each of the 4 windings consists of 200 turns of No. 34 enamelled wire. The core used in

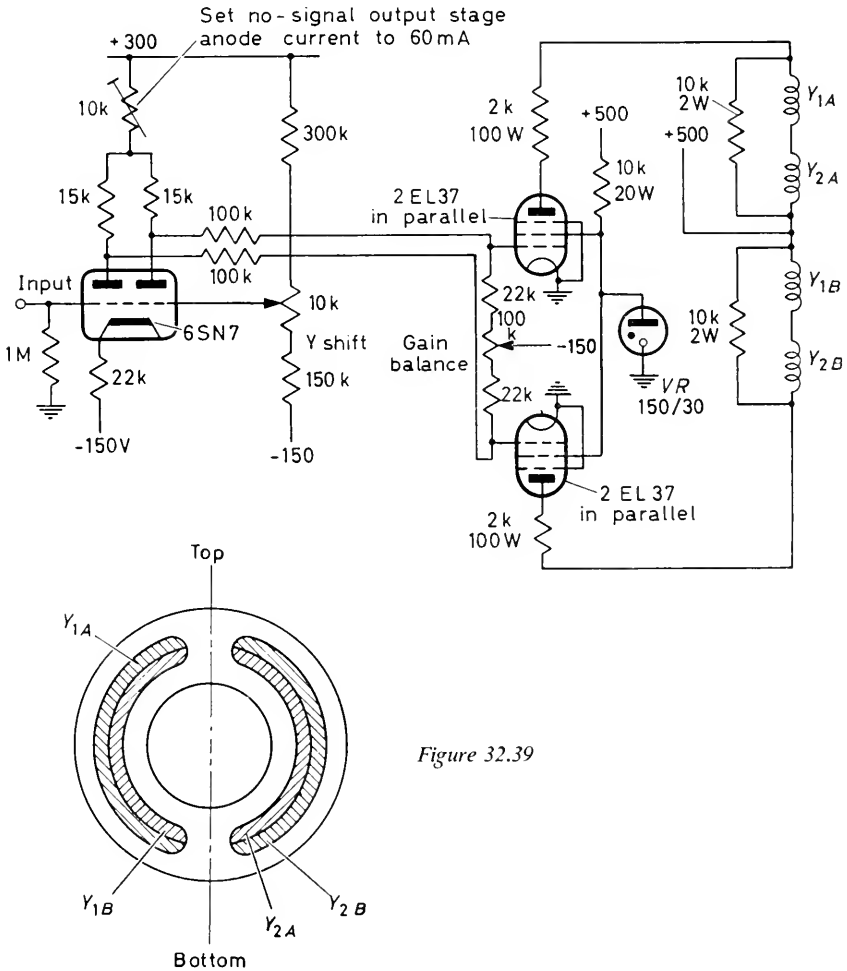


Figure 32.39

the deflector coil assembly comprises two Ferroxcube mouldings, part FX 1154, spaced $\frac{1}{2}$ in.

Circuits for supplying time-base deflector coils have been discussed in Part I.

Electronic brightness control of cathode ray tubes

When photographing a triggered time base it is helpful if the electron beam can be suppressed automatically between sweeps; otherwise there is a risk of fogging of the record by the relatively bright resting spot. It was pointed out in Part I that the triangular wave generators, such as would be

VISUAL INDICATORS

used for time-base waveform production, generate also a positive-going square wave at some point in the circuit. If this wave can be coupled to the grid of the cathode ray tube the required beam suppression is achieved; the brightness control is set so that only during the forward stroke of the time base is the grid bias of the cathode ray tube sufficiently reduced for beam current to flow.

The coupling between the source of the square wave and the C.R.T. grid requires comment. If a magnetic tube is used, the $-ve$ side of the EHT supply is usually earthed and it is not too difficult to devise a direct coupling

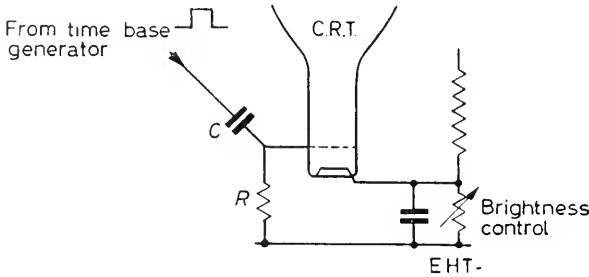


Figure 32.40

scheme employing a neon lamp or potential divider. With electrostatic tubes, the $+ve$ side of the EHT supply is at or near earth potential, and the C.R.T. grid is about 2 kV below earth.

A possible method is to employ resistance-capacitance coupling, remembering that the capacitor must be rated to withstand the EHT voltage (Figure 32.40). The technique is open to two serious objections:

(1) The leakage resistance of C must be much greater than the resistance R , otherwise it will not be possible to suppress the beam. This may restrict

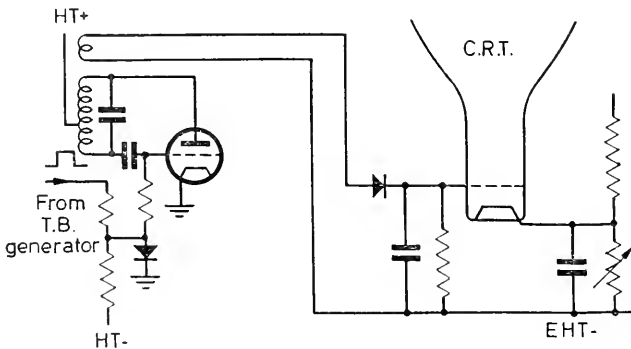


Figure 32.41

the upper possible CR product to a time constant insufficient to pass the brightening-up waves accompanying slow sweeps. It is difficult to brighten up a trace of duration greater than about 100 msec by this method.

(2) The ripple on the EHT supply is now applied to the cathode of the C.R.T., but not to the grid. If the EHT is generated by rectification from

CATHODE RAY TUBE

50~, the sweep will undergo severe brightness modulation at 50~ unless the supply is exceptionally well smoothed.

A better method of achieving trace-brightening is to employ an L-C oscillator, which may conveniently operate at radio-frequency, which is switched on by the square wave from the time-base generator. The oscillations are picked off inductively (*Figure 32.41*) or capacitively (*Figure 32.42*) and rectified to produce a positive-going square wave at the grid of the cathode ray tube. The oscillator may conveniently be switched by returning

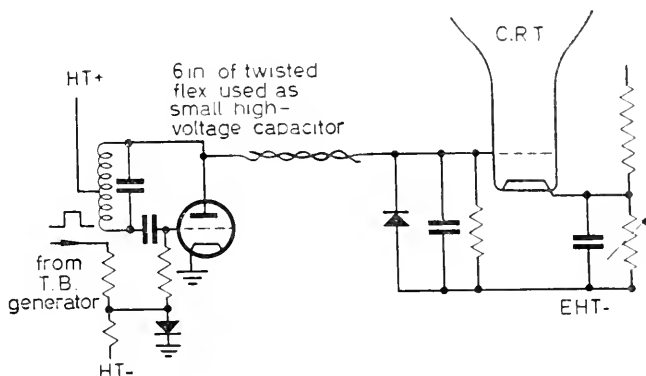


Figure 32.42

the grid leak to a point which is earthy during the forward strokes of the time base, but negative to earth (so that the valve is cut off) at other times.

Hybrid cathode ray tubes

These are a recent development in which the deflection is electromagnetic and the focusing is electrostatic. This retains the advantage of magnetic tubes, in that they are short in comparison with their screen size, and that good focus and brightness are possible; the design is such that the requisite focusing anode potential with respect to the cathode is quite low, about 300 V, enabling both the focus and brightness potentials to be derived from potentiometers across the HT supply, and a negative-earthed EHT supply is used across which no resistor chain is necessary. Hence the drain on the EHT generator is low.

Electrostatic focusing has an advantage over permanent magnet in that remote adjustment does not require complicated mechanical linkages, and over electromagnetic in that almost no power is consumed. It is certainly lighter and it is claimed that it is cheaper.

Multi-channel cathode ray tube displays

It is often valuable to be able to display two varying quantities simultaneously, and we have discussed one method of doing this—the Cossor double-beam tube. Another is to have a number of electrode systems in one glass envelope, all projecting on to one screen; this scheme is feasible with electrostatic tubes but not with magnetic. Two-channel and four-channel tubes may be bought from stock, and eight-channel models to special order

VISUAL INDICATORS

(20th Century Electronics, Ltd.). Multi-gun tubes are expensive, however, and there are two other possibilities.

Beam-switching—A single gun tube is used, and the *Y* deflection system is connected serially to a number of input quantities. If the switching rate is fast enough it appears as if there were a number of distinct traces on the screen. Woods¹⁰ and Wood and Keenan¹¹ may be consulted.

Voltage coincidence—This has been discussed by Reeves¹² and the author¹³. The C.R.T. spot is made to scan rapidly over the entire height of the screen. The alternating voltage which causes this is also applied to a bank of circuits in which it is compared with the signals to be displayed. Whenever the scan voltage coincides with the instantaneous value of one of the signal voltages, the spot is momentarily brightened. In reference 13 the whole question of multi-channel presentation is discussed, and an attempt has been made to show that the voltage-coincidence method scores heavily in cheapness for a small number of channels, and in frequency response for a large number.

REFERENCES

- ¹ STRAFFORD, F. R. W. *Wireless World* 59 (1953) 90
- ² SCROGGIE, M. G. *Radio Laboratory Handbook* London; Iliffe
- ³ DE GRUCHY, J. *Wireless World* 59 (1953) 425
- ⁴ JOHNSON, G. *Wireless World* 61 (1955) 31
- ⁵ COLLINSON, J. D. *Wireless World* 61 (1955) 428
- ⁶ DAVIES, J. R. *Wireless World* 61 (1955) 297
- ⁷ FLEMING WILLIAMS, B. C. *Wireless Engineer* 17 (1940) 61
- ⁸ MOULLIN, E. B. *The Principles of Electromagnetism* Oxford; Clarendon Press
- ⁹ COCKING, W. T. *Television Receiving Equipment* London; Iliffe
- ¹⁰ WOODS, R. W. *Electronics* 28 (1955) 135
- ¹¹ WOOD, K. E. and KEENAN, T. C. *Electronic Engineering* 28 (1956) 105
- ¹² REEVES, R. J. D. *Wireless World* 62 (1956) 85
- ¹³ DONALDSON, P. E. K. *Electronic Engineering* 29 (1957) 78

TRANSDUCERS

K. E. MACHIN

INTRODUCTION

The use of electronic instruments is not confined to purely electrical measurements. By means of suitable devices mechanical displacements, light intensities, temperatures and other physical quantities may be converted into electrical signals. When these signals have been amplified or modified by electronic circuits they are usually re-converted to displacements, either of a meter pointer or of a cathode ray tube spot. All the devices which convert quantities of one type into equivalent quantities of another type are known as transducers; thus a mechano-electrical transducer produces an electrical output corresponding to a mechanical input, while an electroacoustical transducer converts an electrical signal into the equivalent sound.

Thermoelectrical and opticoelectrical transducers are discussed in Chapters 29 and 28 under their more familiar names of electrical thermometers and photocells; the cathode ray tube, yet another example of a transducer, is described in Chapter 32. The present chapter deals only with mechano-electrical (m-e) and electromechanical (e-m) transducers, and with their application to biological research.

Electronic instrumentation inevitably involves expense and complication, and it is therefore important to consider under what circumstances its use is preferable to visual or mechanical methods of achieving the same end. Is it better, for example, to measure the contraction of a muscle with a lever, a stylus and a rotating smoked drum, or with a mechano-electrical transducer, an amplifier, a cathode ray oscillograph and a moving film camera?

Electrical transducer techniques are particularly valuable when the mechanical quantities to be measured either vary very rapidly or are very small. A mechanical recording system inevitably possesses considerable inertia, and it will therefore be unable to follow rapid variations of input. A transducer, on the other hand, can be very small and light, and be made to reproduce faithfully practically any movement which may originate from a biological preparation. When very small mechanical quantities are to be measured, refined engineering techniques are needed to produce the necessary mechanical amplification. Electrical amplification of the signals from a transducer, on the other hand, can be provided straightforwardly and to an almost unlimited extent. Furthermore, a mechanical recording system, particularly one of large amplification, will produce considerable 'loading' of any preparation to which it is attached; this effect can be made very small by using instead a suitable transducer, and moreover is quite independent of the amount of subsequent amplification.

TRANSDUCERS

By the use of transducers and electronic instruments it is often possible to present information in a form more convenient for subsequent analysis. Thus the simultaneous variation of several quantities may be recorded by a multi-channel pen recorder, even though the quantities may be as diverse as pressure, acceleration, volume and voltage. Furthermore, it is possible by suitable circuit techniques to perform mathematical operations before the final recording; if, for instance, information about the rate of change of a mechanical quantity is required, it can be evaluated electronically and recorded directly. Considerable time may thus be saved in the subsequent record analysis.

Finally, mention should be made of the possibilities of telemetry, whereby measurements made with transducers on a freely moving animal can be transmitted via a radio link to a remote recorder^{1,2}.

Throughout this chapter it will be emphasized that the transducer must be chosen carefully to suit the experiment. For this reason there can be no single 'general-purpose' transducer, or even one class of transducer suitable for all biological applications. Very few commercial transducers are available; fewer still are useful in biology. Inevitably, then, the biologist may be faced with the necessity of constructing transducers to fit his own experiments. Very few dimensioned drawings of transducers are given here; hardly any photographs are used; instead, an attempt has been made to explain the basic principles of the various types of transducer, and to illustrate some typical designs in the hope that the reader will then be able to choose the best type of transducer for his particular application, and, if necessary, design and construct it himself.

ANALYSIS OF THE MECHANICAL PROPERTIES OF TRANSDUCERS

In all types of instrumentation it is essential to take into account the limitations of the measuring equipment, and the distortions it may impose on the measured quantity. Thus a lever attached to a biological preparation will have a natural frequency of oscillation, and for faithful reproduction the input quantity must vary slowly compared with the natural period. Furthermore, if a very rapid change of input is applied to such a lever the recorded output will be a damped oscillation bearing no relationship whatsoever to the input quantity. Although the mechanical components of a transducer are usually small and light, they too will have a resonant frequency, which sets a limit to the transducer's speed of response to a changing input. Since transducers are particularly useful in dealing with quantities varying too rapidly for mechanical recording this limit may still be a significant one.

The formal technique for analysing the response of a mechanical system to a given input involves the solution of the differential equation of the system. However, electrical engineers, faced with the same problem in circuit analysis, have evolved rapid geometrical and algebraic techniques for avoiding the necessity for such a solution; in this section it will be shown that circuit analysis techniques can be applied with but little modification to the analysis of mechanical systems. Thus the theory and results of Chapter 5 can be used

directly to evaluate the performance of a mechanical device, whether its final output be mechanical or electrical.

Electrical circuits are compounded of three principal components: inductance, resistance and capacitance. In a similar way, we may analyse mechanical devices into components of mass, resistance and compliance. Every part of a mechanical device possesses mass, and will require a force to accelerate it. This force F_m is given by

$$F_m = m \frac{d^2x}{dt^2}$$

where m is the mass, x is the displacement and t is time. Viscous resistance produces a force proportional to the velocity of movement:

$$F_R = R \frac{dx}{dt}$$

where R expresses the magnitude of the viscous effects. It is usual to consider resistance in mechanical circuits as purely viscous, not frictional, for two reasons. First, if resistance is required in a mechanical device it is usually more convenient to employ some form of fluid damping which is more controllable than solid friction. Secondly, solid friction gives a maximum force while the body is still *stationary*, this force falling off as movement takes place; the mathematical expression of this effect is intractable. In any well-constructed device it is possible to reduce solid friction to small proportions by lubrication, and the assumption of purely viscous resistance seldom leads to great inaccuracy.

Compliance is the property of springs; it is the reciprocal of the more familiar 'stiffness'. The relationship between force and displacement thus becomes

$$F_C = \frac{1}{C} x$$

where C is the compliance of the spring. The units in which the mechanical components are measured are therefore dynes per cm/sec² (or grammes) for mass, dynes per cm/sec for resistance and cm per dyne for compliance.

Mechanical components, like electrical ones, are never 'pure'; a spring possesses some mass, and even the oil in a damping dashpot is slightly elastic. However, as in electrical circuits, one of the components is generally predominant, although neglecting the secondary property of a component may occasionally lead to trouble (*cf.* page 165).

Inspection of the three equations given above shows that they are similar to the equations

$$V_L = L \frac{d^2q}{dt^2} \quad V_R = R \frac{dq}{dt} \quad V_C = \frac{1}{C} q$$

given in Chapters 2, 3 and 4 for the voltages across inductance, resistance and capacitance respectively. It is upon this similarity that the use of electro-mechanical analogies is based. The differential equations of a mechanical system are identical to those of an electrical system if for force we write

TRANSDUCERS

voltage, for displacement charge, while the three components, mass, resistance and compliance become inductance, resistance and capacitance. Were this the final step there would be little point in the technique, as we should have the same differential equation to solve with different symbols. However, the fact that the equivalent equations can be set up means that the techniques of Chapters 3, 4 and 5 (vector diagrams and the use of the operator j) become available for the solution of mechanical problems. Any mechanical 'circuit' can now be drawn as an equivalent electrical circuit. The relationship between the forces and displacements in the parts of the mechanical circuit are exactly similar to the relationship between voltages and charges in the corresponding electrical circuit.

It is not usual in electrical circuit analysis to evaluate charge: its rate of change with time, the current, is more commonly derived:

$$i = \frac{dq}{dt}$$

The solution of the electrical equivalent circuit will then give not displacement, but velocity:

$$v = \frac{dx}{dt}$$

Failure to realize this is the first common pitfall of electromechanical analysis; in biological work the displacement is likely to be of greatest interest, and must be derived from the equation

$$x = \int v dt$$

The second pitfall may occur in the transformation of a mechanical circuit into an equivalent electrical one; it is often not self-evident whether certain components must appear in series or in parallel in the equivalent circuit. For this there is a simple rule: mechanical components subjected to the same *force* are equivalent to electrical components with the same *voltage*, i.e. in parallel. Similarly mechanical components experiencing the same *displacement* or *velocity* appear in series in the equivalent circuit.

So far linear translational motion only has been discussed. Similar considerations apply to circular motion, and the equivalent quantities for both types of motion are summarized in *Table 1*.

TABLE 1

<i>Linear motion</i>	<i>Circular motion</i>	<i>Electrical equivalent</i>
Force, dynes Displacement, cm Velocity, cm/sec Mass, grammes Resistance, dynes per cm/sec Compliance, cm per dyne Components subjected to same force Components subjected to same displacement	Couple, dyne-cm Angular displacement, radians Angular velocity, rad/sec Moment of inertia, grammes-cm ² Resistance, dyne-cm per rad/sec Torsional compliance, radians per dyne-cm Components subjected to same couple Components subjected to same angular displacement	Voltage, volts Charge, coulombs Current, amps Inductance, henries Resistance, ohms Capacitance, farads Components in parallel Components in series

As an example of the use of electrical equivalent circuits, the performance of the force-recording apparatus of *Figure 33.1* will be analysed.

MECHANICAL PROPERTIES OF TRANSDUCERS

A pen is constrained to move linearly by a slide of mass m . A spring of compliance C controls the movement of the slide, and the force to be measured is applied through a tension wire. The slide is lubricated and is subject to a viscous resistance R . The simple approach gives the linear movement of the slide as proportional to the force applied.

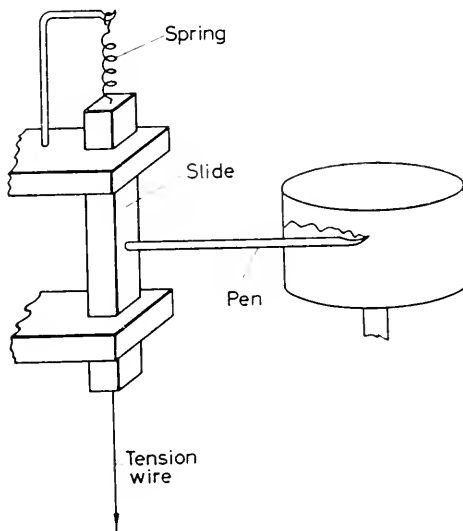


Figure 33.1 Force-recording apparatus

In drawing the electrical equivalent circuit it will be seen that all three components, mass, resistance and compliance, have the same displacement, and that the equivalent circuit thus has inductance, resistance and capacitance in series (*Figure 33.2*). The applied force is the applied voltage; the velocity of the pen is the current in the circuit.

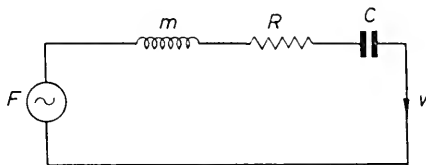


Figure 33.2 Equivalent circuit of force-recording apparatus

Since the pen's displacement, not its velocity, is required, it is necessary to use one of the equations:

$$x = \int v \, dt, \quad q = \int i \, dt$$

A useful artifice to avoid this integration uses the fact that the voltage across the condenser is given by the equation

$$V_C = \frac{1}{C} q$$

TRANSDUCERS

Therefore the voltage on the capacitance represents $1/C$ times the charge; using this result the equivalent circuit may be drawn as in *Figure 33.3*.

This circuit has already been analysed in Chapter 5; it was shown (*Graph 23*) that for a constant sinusoidal input the output is constant at low frequencies, reaches a maximum the height of which depends on the resistance, then falls steeply away at high frequencies. The response of the circuit to a sudden transient input has also been discussed, and is shown in *Graph 22*. From this analysis it is clear that the apparatus will record faithfully periodic forces only if their frequency is well below the resonant frequency. In addition, sudden changes of force will be recorded at best only after a delay of about half the natural period of the apparatus, or, if the damping is low, with many 'overshoots'.

Typical values of the mechanical constants are $m = 10$ g, $C = 1$ cm per kg wt (i.e. 10^{-6} cm per dyne). The resonant frequency would then be about 50 c/s. If the value of the resistance is appropriate for 'critical damping' (see Chapter 5) the response of the apparatus to periodic forces will be only 90 per cent of its expected value at about 20 c/s; the pen will also take about 10 msec to register a sudden change of force.

In evaluating the response of a transducer it is necessary to take into account not only the mechanical properties of the transducer itself but also those of the devices coupling it to the preparation, and in some cases the properties of the preparation itself. For example, an RCA mechano-electronic transducer (described on page 490) has been used by Dr. G. M. Hughes to measure the propulsive force of the jet efflux from a dragon-fly nymph; the arrangement is shown in *Figure 33.4*. Here it is necessary to consider not only the inertia and compliance of the transducer but also the compliance of the support arm, the mass of the insect and the drag of the water on it. *Figure 33.5* shows the equivalent circuit. Here R is the resistance due to water drag, M_1 and M_2 the moments of inertia of the insect and transducer plate respectively, and C_1 and C_2 are the compliances of the support arm and the transducer diaphragm respectively. The input to the circuit is a couple (the product of the required force and the support length), while the output is $1/C_2$ times the angular displacement of the transducer plate; the electrical output of the transducer is proportional to this displacement.

It is often possible to simplify an equivalent circuit, since the component parts are usually of widely differing magnitudes. In the present case, $M_2 \ll M_1$ and $C_1 \ll C_2$. The circuit reduces to that of *Figure 33.6*, which is similar to that of *Figure 33.3*. Now, however, the resonant frequency and damping of the system depend directly on the properties of the biological preparation under investigation. Although the natural frequency of the transducer alone is 12 kc/s, with the insect and support in place it is reduced to 120 c/s, with a damping ratio of about 0.4. In Dr. Hughes' experiments this lowering of the resonant frequency was of no importance, since the jet force varied relatively slowly. However, the analysis of this mechanical system demonstrates the care needed in designing attachments to preparations if the advantages of transducers (sensitivity and wide frequency range) are to be exploited to the full.

The equivalent circuit technique will be used again on pages 492, 495, where it will be shown that the mechanical and electrical properties of an

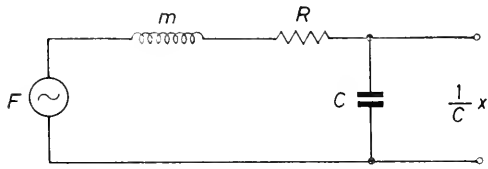


Figure 33.3 Alternative equivalent circuit of force-recording apparatus

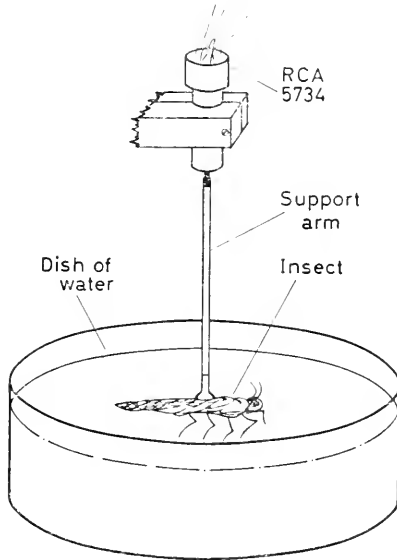


Figure 33.4 Apparatus to measure jet force of dragon-fly nymph

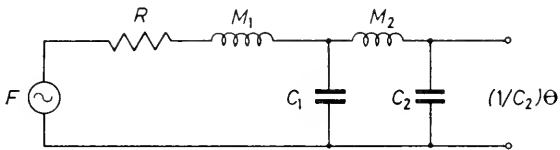


Figure 33.5 Equivalent circuit of the apparatus in Figure 33.4

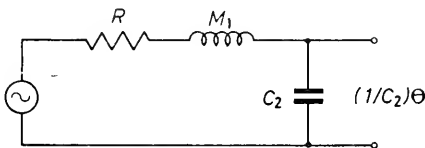


Figure 33.6 Simplified equivalent circuit of the apparatus in Figure 33.4

e-m transducer interact so that the mechanical properties of the transducer can be influenced by the nature of the electrical circuits driving it.

MECHANICAL IMPEDANCE OF TRANSDUCERS

The input impedance of an electrical instrument is a measure of the extent to which the instrument disturbs the circuit under test. When measuring voltage a high-impedance meter is needed; the measurement of current requires a low-impedance instrument (see Chapters 32 and 43). Similarly, the disturbance due to a transducer is assessed in terms of its *mechanical input impedance*; whether this impedance should be low or high depends on the nature of the quantity to be measured. A transducer measuring displacement should require as little force as possible to deflect it, i.e. it must present a 'low impedance'. A force-measuring transducer, on the other hand, should be of high impedance, as it must 'give' as little as possible under the influence of the applied force. Although the impedance of a transducer is given directly by the input of its equivalent electrical circuit, the compliance components usually predominate within the working range, and hence they largely determine the impedance.

Most m-e transducers produce an output dependent upon their displacement; whether they are most suitable for measuring force or movement depends on their impedance. A low-impedance transducer can be adapted to measure force by putting a stiff spring across its input, thereby raising its impedance, while a high-impedance transducer connected to a preparation through a weak spring may be suitable for measuring movement.

It is therefore not intended in this chapter to separate transducers artificially into 'force' and 'displacement' types since these differ only in their impedance; both fundamentally are sensitive to displacement.

The characteristics of e-m transducers can similarly be described in terms of their mechanical output impedance. A low-impedance device is one producing constant force regardless of movement, while a high-impedance device gives a displacement virtually independent of load. Most e-m transducers are fundamentally of low impedance, but if combined with a stiff spring can be made to present a high impedance.

The transformation of a low impedance transducer into a high-impedance one, and vice versa, by the use of a spring is not without its pitfalls. To demonstrate this, and to give further illustrations of mechanical circuit analogies, the equivalent circuits of various transducer arrangements will be considered.

If a very high-impedance m-e transducer, such as might be suitable for measuring force, is connected to a moving preparation through a weak spring so as to measure the displacement of the preparation, the equivalent circuit will be as in *Figure 33.7*.

Here m , c are the masses and compliances of the short elements of length of the spring; C_1 is the (very low) compliance of the transducer. At low frequencies the effect of the masses becomes negligible and the circuit reduces to the compliance of the spring as a whole in parallel with C_1 . However, at high frequencies, the spring with its distributed mass and compliance behaves like an electrical transmission line, with distributed inductance and

MECHANICAL IMPEDANCE OF TRANSDUCERS

capacitance^{3,4}. The detailed properties of such a system are outside the scope of the present analysis, but various salient features should be noted: (1) mechanical signals propagate along the spring with a finite velocity; (2) a phase shift proportional to frequency is thereby introduced; (3) at certain frequencies the signals reflected at the transducer and transmitted back to the input end can cause very severe loading of the preparation. As a result the input impedance of the spring may become very high.

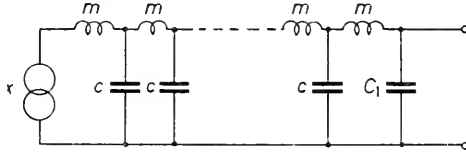


Figure 33.7 *Equivalent circuit of force transducer fed through a weak spring*

To avoid all these effects the time of transmission of a mechanical signal along the spring must be short compared with the periodic time of the signal. Since the velocity of propagation v along a spring is given by

$$v = \frac{1}{(m_1 c_1)^{1/2}}$$

where m_1 and c_1 are the mass and compliance per unit length, it follows readily that the time of transmission along the spring is

$$t = (MC)^{1/2}$$

where M and C are the mass and compliance of the whole spring. It also follows that the frequency at which the input impedance of the spring becomes high is given by

$$f_r = \frac{1}{2(MC)^{1/2}}$$

For example, a spring designed to give a force of 10^4 dynes (10 g wt) per cm of movement might weigh 50 mg, and be 2 cm long. The propagation time along it would then be about 2 ms. Thus sinusoidal mechanical inputs with a frequency of 120 c/s would suffer a phase shift of about 90 degrees; at a frequency of 250 c/s the input end of the spring would appear almost completely rigid.

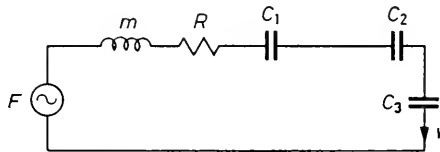


Figure 33.8 *Equivalent circuit for displacement transducer with added stiff spring*

As mentioned above, a low-impedance transducer can be used to measure force by putting a stiff spring across it: the input member of the transducer is then rigidly attached to the preparation. The equivalent circuit is shown in Figure 33.8. Here m , R , C_1 are the mass, resistance and compliance

DYNAMIC MECHANOELECTRICAL TRANSDUCERS

the transducer mechanisms generally differ only in their geometrical arrangements in the two cases. Indeed the input of many so-called linear displacement transducers is, in fact, applied at the end of a rotating arm; the angle over which this arm moves is arranged to be small so that its end moves in an almost straight line.

Transducers for continuous rotary motion, such as generators and motors, are not fundamentally different, but for convenience will be considered in a separate section.

DYNAMIC MECHANOELECTRICAL TRANSDUCERS

As explained above, dynamic transducers do not give a steady output for a steady deflection: in other words, their response does not extend down to zero frequency. It is important, however, to discuss the exact form of the frequency response of a dynamic transducer, and to examine how the output for a constant input displacement falls off as the input frequency approaches zero. This behaviour is different for the two major types of dynamic transducer.

Piezo-electric transducers

Certain crystalline materials, when subjected to mechanical stress, develop surface electric charges. Thus, if a stress is applied to a piece of such material provided with suitable electrodes, charges will appear on the electrodes, and hence a voltage will be developed across them. This is the principle on which piezo-electric transducers operate; the disposition of the electrodes, and the mode of application of the stress, will vary with the material and configuration of the transducer.

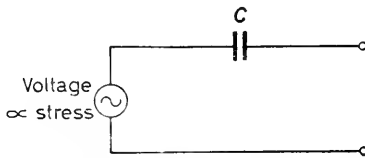


Figure 33.9 Equivalent electrical circuit of piezo-electric transducer

Viewed from the output terminals, a piezo-electric transducer behaves like a fixed capacitor carrying a charge (and hence a voltage) proportional to the applied stress. Were it possible to connect the transducer to an indicating device of infinite resistance (such as an electrostatic voltmeter), the indicator would deflect when the stress was applied, remain deflected so long as it was present, and return to its original zero when the stress was removed. The frequency response of the system would then extend down to zero frequency. However, in practice the transducer is likely to feed an amplifier or other device with finite input resistance; the charges developed as the stress is applied will gradually leak away through this resistance. So an indicator on the output of the amplifier will deflect as the stress is applied, and then gradually return to zero with a time constant of CR , where C is the capacitance of the transducer and R is the input resistance of the amplifier.

All the characteristics of the transducer-amplifier combination, such as steady-state frequency response and transient response, can be determined by giving the transducer the equivalent circuit of Figure 33.9.

TRANSDUCERS

From the results of Chapter 3 it is clear that the transducer-amplifier combination will reproduce faithfully sinusoidal variations of stress only if they occur at a sufficiently high frequency, and that transient variations will be accurately delineated only if they occur rapidly enough. *Graph 9* shows the steady-state behaviour of a piezo-electric transducer; the transient response will be similar to *Graph 7*.

Typical piezo-electric transducers may have capacitances of 100–2,000 pF; accurate reproduction of frequencies down to a few c/s will thus involve an amplifier with an input resistance of many megohms. A cathode follower (Chapter 11) using if necessary an electrometer valve (Chapter 24) provides a suitable high input-resistance first stage for this amplifier.

Two piezo-electric materials are suitable for biological transducers: Rochelle salt and barium titanate. Rochelle salt is usually assembled into a double slab (or 'bimorph') and the applied stress may be in the form of a torque ('twister' bimorph) or a flexure ('bender' bimorph). Bimorphs are used extensively in gramophone pick-ups; the insert cartridge from such a pick-up forms a convenient transducer, stress being applied to the needle through a tension wire. As a transducer such a device has a high mechanical impedance and behaves like a stiff spring of compliance about 10^{-6} cm per dyne (10μ per g wt). As mentioned on page 478, such a transducer is immediately applicable to the measurement of force, but has to be connected through a light spring when used to measure displacement. Typical gramophone pick-up cartridges have a capacity of 500 pF, and give an output voltage of a few mV per dyne, i.e. several volts for a force of 1 g wt.

Barium titanate bars are in common use as accelerometers in vibration detection. Here the stress in the material is generated by inertial forces as the transducer is subjected to acceleration. Some of these accelerometers are usable as transducers if a tension wire is attached to an appropriate point. A typical barium titanate transducer, used in tension, has a compliance of about 10^{-10} cm per dyne (1μ per kg wt). This high impedance makes it immediately applicable to the measurement of force. An output of the order of $20 \mu\text{V}$ per dyne (20 mV per g wt) can be obtained in this way.

All piezo-electric transducers reach an upper frequency limit when their mechanical resonant frequency is approached. This frequency is determined by the compliance of the transducer, in combination with the effective mass of its moving parts plus the mass of the preparation and attachments. With reasonable care resonant frequencies of many kc/s can be achieved.

In biological research piezo-electric transducers are likely to be of the greatest value for the measurement of rapidly changing forces, such as the contraction of fast muscles. Here the fact that the response does not extend down to zero frequency is little disadvantage, while the small size, high upper frequency limit, large output and low compliance can be particularly useful.

Two typical piezo-electric transducers are shown in *Figure 33.10*.

Electrodynamic transducers

The principle of all electrodynamic transducers is as follows. The mechanical input is arranged to vary the magnetic field through a coil; this variation induces a voltage in the coil. The voltage is proportional to the rate of

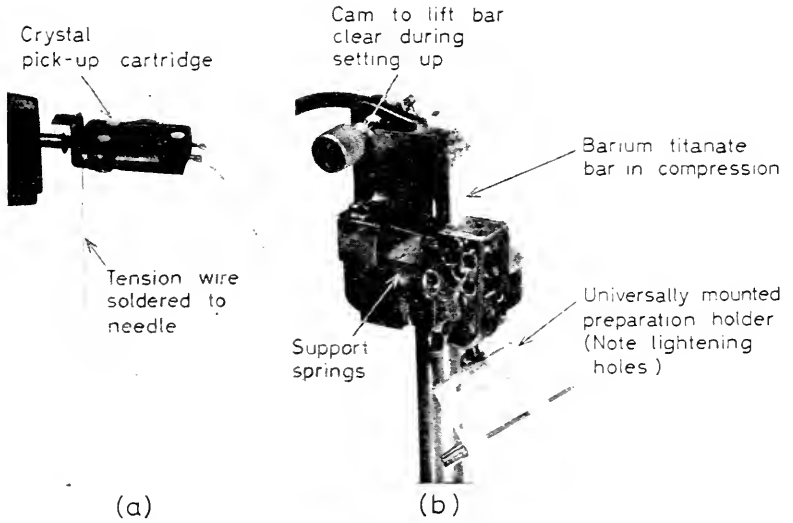


Figure 33.10 Piezo-electric transducers: (a) gramophone pick-up; (b) barium titanate bar

change of magnetic field, and therefore with all common configurations to the velocity of the mechanical input.

If a sinusoidal mechanical input is applied to such a transducer, the resulting peak velocity will be proportional both to the amplitude of the displacement and to its frequency. Thus for a constant amplitude input to an electrodynamic transducer, the voltage output will be proportional to frequency; this characteristic is shown in *Figure 33.11*. Comparison with

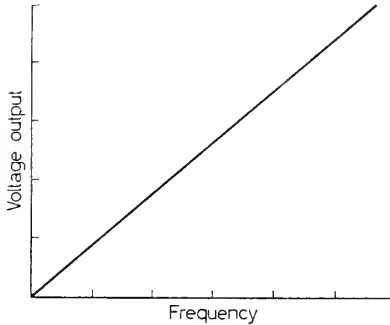


Figure 33.11 Response of electrodynamic transducer for constant input displacement

Graph 9 shows that there is an essential difference between electrodynamic and piezo-electric transducers, although both give no output at zero frequency.

Electrodynamic transducers can be constructed with very low mechanical impedance, suitable for use for measuring displacement. However, the rising frequency characteristic implies that the electrical output will not give a true picture of the displacement. Although this frequency characteristic may be linearized by passing the output through an integrating circuit (Chapter 3), it is usually possible to choose a more satisfactory type of transducer for the purpose. However, if information about the velocity of a system is required, the output of an electrodynamic transducer, suitably amplified, can be used to record this quantity directly. This avoids subsequent graphical differentiation of a displacement record and is the most useful application of this type of transducer.

For qualitative work when, for example, only the frequency of a vibration or the time of occurrence of a transient mechanical disturbance is required, the simplicity and high sensitivity of certain types of electrodynamic transducer are valuable. Thus, if an insect is attached to the armature of a moving-iron transducer (e.g. a magnetic gramophone pick-up), the frequency of the output will correspond to the wing-beat frequency, even though the output waveform gives little information about the forces to which the insect is subjected. Again, a piece of iron mounted on a rotating device will give several volts output in a nearby coil; this output may be used, for example, to synchronize a photographic flash with the rotation.

Moving-iron transducers—In this type of transducer the coil is fixed, and a moving armature varies the reluctance of a magnetic circuit (Chapter 4) passing through it. The m.m.f. is provided by a permanent magnet, which

TRANSDUCERS

may either remain fixed with the coil assembly or may be incorporated in the moving armature.

Many configurations are possible: two common ones are shown in *Figure 33.12*. The compliance of the transducer is dictated entirely by the

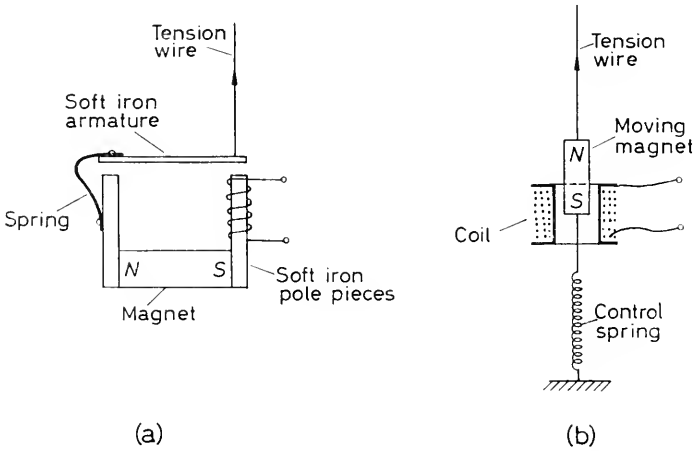


Figure 33.12 Two moving-iron transducers shown diagrammatically

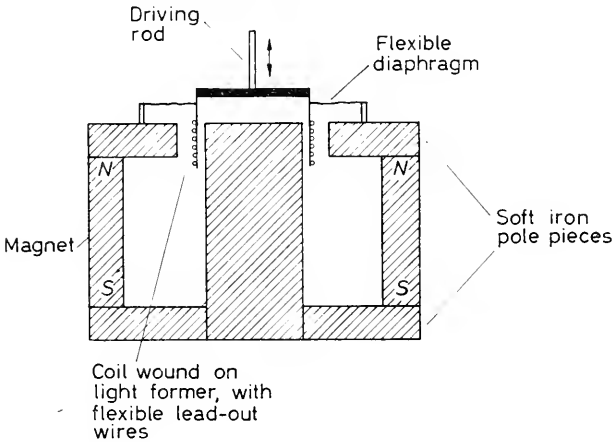


Figure 33.13 Section of moving-coil transducer

spring on which the armature is mounted, and it may be chosen to suit the application. It is not possible to give any figures for output voltage, as it depends upon the number of turns on the coil, the strength of the magnet and on the geometry of the armature and coil assembly.

The relationship between the output voltage and input velocity with a moving-iron transducer is inherently non-linear. Although the non-linear effects can be minimized by operating the transducer over only a small range of displacement, moving-iron transducers are best regarded as only suitable for qualitative work.

Moving-coil transducers—If a coil is moved through a uniform stationary

STATIC MECHANOELECTRICAL TRANSDUCERS

magnetic field, the output voltage is accurately proportional to the velocity of the coil. Such a system is illustrated in *Figure 33.13*; it will be seen that it is similar to the ordinary moving-coil loudspeaker.

With a moving system weighing several grammes and a suspension compliance of about 10^{-6} cm per dyne (10μ per g wt), the usual moving-coil transducer has a somewhat inconvenient mechanical impedance. When used for measuring velocity its compliance is so low that undue loading of the preparation will occur, while as a force transducer its resonant frequency is uncomfortably low. However, when loading effects are of little importance a moving-coil transducer provides an accurate measurement of velocity with an output which may be as much as 1 V per cm/sec. Transducers of this type can readily be adapted from commercial moving-coil loudspeakers.

STATIC MECHANOELECTRICAL TRANSDUCERS

Dynamic transducers, being self-generating, derive all their electrical output power from the mechanical system under observation. Static transducers, however, require an external source of electrical power; they control mechanically how much of this power is delivered at the output. Theoretically they can therefore cause no loading of the mechanical system whilst delivering an indefinitely large electrical output. For this reason they are more generally useful than dynamic transducers.

The circuit element which is mechanically controlled may be resistance, inductance or capacitance; static transducers will be discussed under these three headings. The special circuits used with some of the transducers will be described in a later section.

Variable resistance transducers

Transducers using variable resistance occur in three forms: (a) wiping contact on resistive track; (b) moving electrode in liquid resistance; (c) deformation-sensitive resistive element. These types have very different characteristics and will be considered separately.

Resistive track transducers—In this type of transducer the wiping contact may move either round a circular resistive track or linearly along a straight track. The circular type is eminently suitable for use with circular motion, but requires levers or a rack when used for linear motion. On the other hand, such devices are readily available in the form of the ordinary radio potentiometer, and, in refined forms, are capable of the highest accuracy of any type of transducer. Linear variable resistances will usually have to be constructed for a particular application, but then may be applied directly to the measurement of linear movement.

For low- and medium-accuracy work, the potentiometers described in Chapter 20 are suitable; the degree of accuracy to be expected is discussed there. It is possible, however, to obtain potentiometers with a linearity rather better than 0.1 per cent; they are known as 'cam-corrected' potentiometers. These devices are basically very high-grade wirewound potentiometers, but with the addition of an adjustable face-cam by which the inevitable small inaccuracies may be corrected at, typically, eight points. This correction procedure, which is of course carried out by the manufacturer,

TRANSDUCERS

makes the potentiometer extremely accurate at the eight points, and also gives a great improvement in the accuracy elsewhere.

All types of potentiometer are available with non-linear 'laws'; thus the resistance may be made to vary as, for example, the square of the shaft rotation. The accuracy with which these functions can be reproduced is usually rather lower than that of the corresponding linear-law potentiometer.

Potentiometers for linear operating movement can be constructed for special applications using carbon tracks, wound cards or even very fine single resistance wires. Their characteristics are dependent entirely upon the precision of workmanship, though such potentiometers are never likely to be as accurate as the corresponding commercial rotary units.

The use of potentiometers as transducers is limited to those occasions when considerable mechanical power is available. The potentiometers described may require a driving torque of 10^4 – 10^5 dyne-cm (10–100 g wt-cm), although units of special design are available with a driving torque of well under 10^3 dyne cm. Furthermore, if complex and inefficient linkages are to be avoided ample movement must be available. Potentiometers require positive drive in both directions, so that in certain cases a return spring may be necessary.

Liquid potentiometers—The disadvantages of large force and movement inherent in track potentiometers may be avoided by using a liquid column as the resistive element. The moving contact is merely an electrode moving in the liquid, and the forces on it are negligible. Minute electrode movements can be used, as the potentiometer can be made on a very small scale.

It is impracticable to construct a linear variable resistance (a 2-terminal device—as distinct from a potentiometer, a 3-terminal device) using a liquid element. Unless the cross-section of the electrodes is the same as that of the tube containing the liquid the variation of resistance with electrode distance is very non-linear and almost impossible to calculate. With a pair of small electrodes almost all the voltage drop occurs within a small distance of the electrode surface, and it will be found that the electrode spacing has but little effect on the resistance between them. On the other hand, with the potentiometer configuration, large fixed electrodes fitting the whole cross-section of the container can be used, with a very small moving electrode acting virtually as a potential-exploring probe. If it is not practicable to use fixed electrodes covering the whole area, a linear voltage/distance relationship is still obtained in the centre, provided that the distance between the electrodes is large compared with their dimensions. In this case, of course, only a fraction of the voltage applied across the fixed electrodes is available as output.

It is desirable in order to avoid polarization of the electrodes to use alternating current with a frequency of at least 250 c/s for energization. Care should also be taken with the nature of the container, electrolyte and electrodes, to avoid chemical action, evaporation or other deterioration. Platinum electrodes are usually safe, while polythene, Perspex and glass are suitably inert materials for the container. Electrolytes which have been used include glycerin, and, for sealed-off transducers, a mixture of hydrochloric acid and alcohol.

The mechanical input impedance of such a transducer is dependent almost

STATIC MECHANOELECTRICAL TRANSDUCERS

entirely on the suspension of the moving electrode. In some cases it is possible to attach this electrode directly to the preparation, in which case the loading effect of the transducer can be made negligible. When the transducer is operated through a thread or chain, some form of spring suspension is necessary; the compliance of this spring determines the impedance of the transducer.

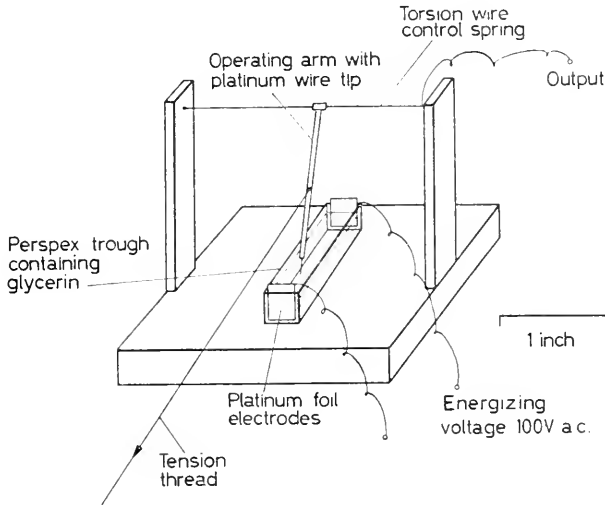


Figure 33.14 Liquid potentiometer for displacement measurement

Two examples of liquid potentiometers will serve to illustrate their possibilities. The first, due to J. A. Popple, is illustrated in Figure 33.14. An output of 50 V per cm movement can be obtained with an input compliance of more than 10^{-2} cm per dyne (10 cm per g wt). This transducer has been used to study the gill movements of fishes. The second design (Figure 33.15) is due to Dr. D. W. Kennard, and has several novel features. First, a very stiff suspension is used as the transducer is designed for the measurement of muscle forces under isometric conditions. Secondly, d.c. energization is used. It is found that, provided the voltage across the electrolyte is low enough (~ 1 V), polarization effects stabilize after about an hour, and that no bubbling occurs. The suspension compliance is about 10^{-7} cm per dyne (1μ per g wt), giving a usable response up to about 5 kc/s. A sensitivity of $1 \mu\text{V}/\text{dyne}$ (1mV per g wt) is obtained.

Resistance transducers based on deformation—If a length of wire is subject to stress the resulting elongation and change of area alters its resistance. This principle is used in the resistance strain gauge, which consists of many turns of resistance wire wound on a paper former. In the engineering applications for which these gauges were designed the unit is cemented to a structure under test; the exceedingly small extensions in the structure can then be measured electrically. Strain gauges can, however, be used unmounted, or bonded on to flexible materials such as rubber. In this way transducers with a very low compliance (10^{-8} cm per dyne, 0.1μ per g wt), suitable for measuring force without any appreciable extension, can be made.

TRANSDUCERS

Unfortunately the maximum change of resistance which can be obtained before the gauge wires pass their elastic limit, and hence change their calibration, is about 4 per cent. Furthermore, the power dissipated in the gauge must not exceed, typically, about 25 mW lest its characteristics be affected. Thus, with an unmounted strain gauge of 1,000 ohms in a suitable bridge, a maximum output of 0.2 V would be obtained for a force of about 1 kg wt. Their sensitivity to humidity, however, severely limits the applications of unmounted strain gauges in biological research.

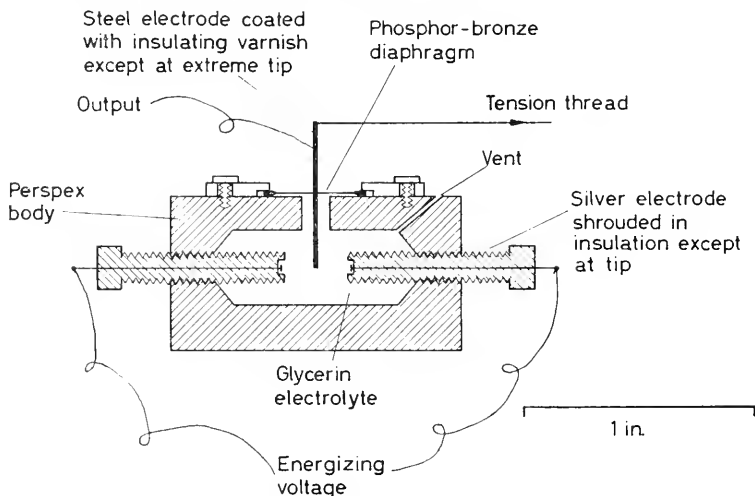


Figure 33.15 Liquid potentiometer for force measurement

Another type of transducer is based on the properties of carbon contacts. If a circuit includes a number of contacts between pieces of carbon, the resistance of the circuit depends on the mechanical pressure at the contacts. Devices based on this effect include the carbon pile, a stack of discs under end-wise pressure, and the carbon microphone, with granules trapped between a fixed and a moving electrode. All such devices tend to be non-linear, noisy and subject to drift, and are therefore best considered for qualitative work only.

Variable inductance transducers

The inductance of an iron-cored inductor may be varied by moving some part of its magnetic circuit mechanically. This is the principle of the operation of the variable inductance transducer.

Like the moving-iron transducer, to which it bears a superficial resemblance, the variable inductance transducer may occur in many different configurations. Some of these are illustrated in Figure 33.16. Unlike moving-iron devices, no permanent magnets are used and magnetic forces are small. Armature control springs may be correspondingly lighter, and in some cases it is possible to attach a light armature directly to the preparation and use no springs at all. Alternatively, by using very stiff armature control springs and very small movements, transducers of a high mechanical impedance suitable for force measurement may be constructed.

TRANSDUCERS

Capacitance transducers have many properties in common with inductance transducers; thus they may easily be constructed so as to have a mechanical impedance suitable for any particular purpose. In addition, they too are fundamentally non-linear and must therefore be operated over a restricted

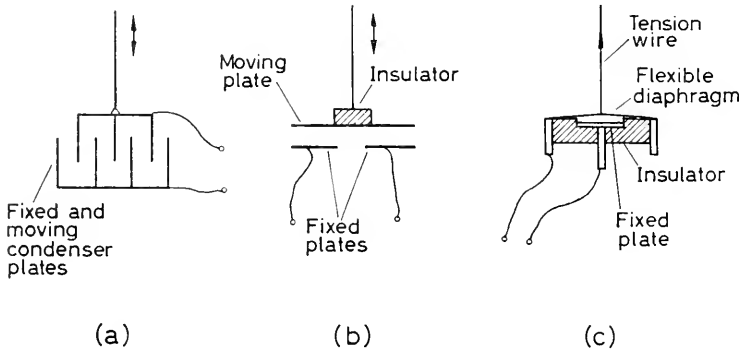


Figure 33.18 Variable capacitance transducers

range if linearity is desired. Transducers of this type are potentially extremely sensitive; with close electrode spacing (and, of course, correspondingly good workmanship) movements of 0.01μ may easily be detected. When the moving electrode is controlled by a stiff spring to give a high mechanical impedance such a transducer can be used to measure force with very little deflection and a high resonant frequency.

Miscellaneous static transducers

In this section two transducers will be described which while basically of the variable resistance type achieve this variation in such an unusual way that they are best classified as miscellaneous.

The mechano-electronic transducer (RCA Type 5734) is illustrated in *Figure 33.19*. The anode of a very small triode valve is flexibly mounted on a metal diaphragm and is continued outside as an operating shaft. Sideways deflection of this shaft varies the anode-cathode spacing, and hence the anode current of the valve.

With the mechanical input applied to the end of the 3 mm long operating shaft, the compliance of the transducer is 10^{-7} cm per dyne (1μ per g wt). This gives a very high mechanical input impedance well suited to the measurement of force. The resonant frequency of the transducer alone is 12 kc/s, although this is inevitably reduced by any mass attached to the operating shaft.

The circuit used is shown in *Figure 33.20*; it will be seen to be a variant of the Wheatstone bridge. The sensitivity of the transducer in this circuit is about 1 mV per dyne (1 V per g wt) applied at the tip of the operating shaft. By using a longer lever on the operating shaft this sensitivity can be increased, but only at the expense of lower input impedance and resonant frequency.

Mechano-electronic transducers are inevitably rather delicate, both mechanically and electrically, and it is essential before such a device is used to consult the makers' literature⁵ to find out the recommended precautions.

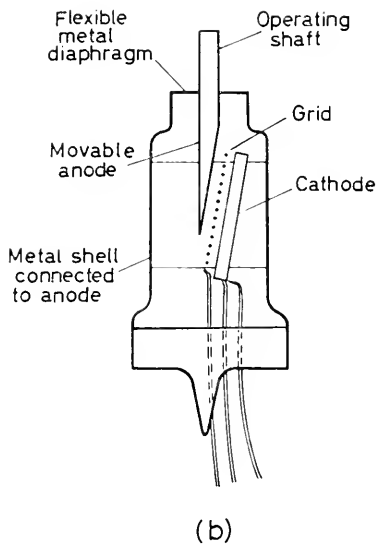
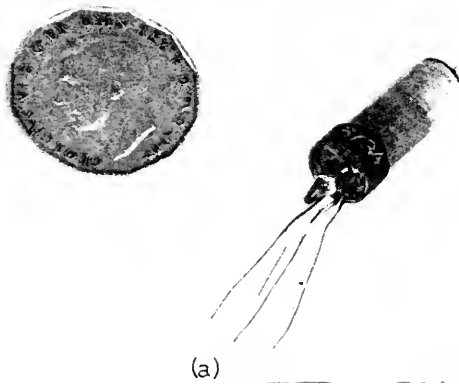


Figure 33.19 RCA 5734 mechano-electronic transducer: (a) external appearance; (b) diagram of internal construction



STATIC MECHANOELECTRICAL TRANSDUCERS

The photoelectric transducer, illustrated diagrammatically in *Figure 33.21*, has become particularly convenient since the introduction of the extremely small phototransistor (Chapter 28). Small transducers of good sensitivity,

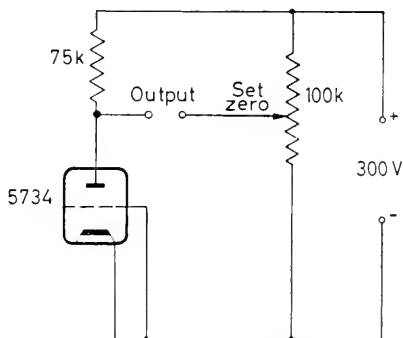


Figure 33.20 Circuit for RCA 5734 transducer

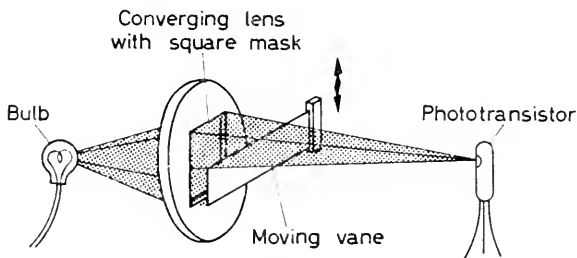


Figure 33.21 Photoelectric transducer

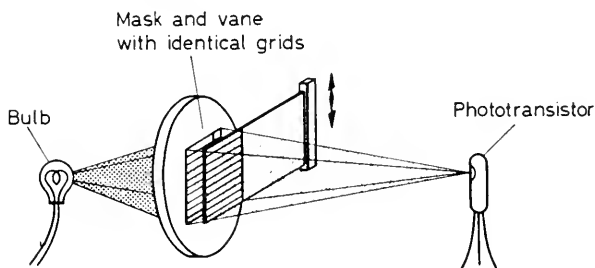


Figure 33.22 Photoelectric transducer for very small displacements

which need no contact with the preparation, can now be constructed. The upper frequency limit of such a transducer is set by: (a) the cut-off of the photocell—3 kc/s for a phototransistor; (b) the resonant frequency of the vane, which by careful design can be made as high as several kc/s.

The sensitivity of a photoelectric transducer using a plain aperture (as illustrated) may be about 50 V/cm. This can be increased many-fold by using a double grid to reduce the apparent aperture; this is shown in *Figure 33.22*.

TRANSDUCERS

With suitably shaped vanes in front of plain apertures it is possible to obtain almost any non-linear relationship between displacement and voltage output.

ELECTROMECHANICAL TRANSDUCERS

All the dynamic m-e transducers described in an earlier section may be used the other way round, i.e. to provide a mechanical output for an electrical input. Since general descriptions have already been given, the present section will deal only with the characteristics which the devices display when used as e-m transducers. Certain configurations suitable only for e-m use will also be mentioned.

Piezo-electric transducers

Although piezo-electric transducers used electromechanically are only capable of minute movements, the large inherent stiffness of the piezo-element results in a very high mechanical output impedance, and the movements can be sustained against large opposing forces.

An interesting use of a piezo-electric transducer is due to Pascoe⁶, who had experienced difficulty in inserting microelectrodes through cell walls without breaking the tip of the electrode. He then mounted the electrode on a 'bender' crystal and was able to jerk it forwards about 20μ by the application of a d.c. voltage to the crystal. This sudden movement effected the penetration without damage to the electrode.

Electrodynamic transducers

Mechanoelectric transducers are usually operated into the input of an amplifier. Thus when an electrodynamic transducer is used mechano-electrically a negligible current is drawn from the transducer. If a current were allowed to flow the transducer would then behave electromechanically and a reaction on the preparation would occur: the mechanical impedance of the transducer would in fact be modified by the electrical impedance into which it worked. This effect does, in fact, occur when an electrodynamic transducer is used electromechanically, as it will be fed from a source of electrical power having a definite impedance. This impedance will modify the mechanical output impedance of the transducer.

To illustrate this effect, consider an ideal electrodynamic transducer consisting of a length of wire of zero mass and resistance, freely situated in a magnetic field. If a constant current generator is connected to the wire it will experience a constant force regardless of any movement it may suffer. On the other hand, if a constant voltage generator is used the wire will move until it generates a back e.m.f. equal to that of the generator, i.e. it will move with constant velocity regardless of load. Thus, a transducer fed from a high-impedance electrical supply has a low mechanical output impedance; a low-impedance drive to the same transducer gives it a high mechanical output impedance. It can be shown that if the electrical supply has a resistive impedance, as is usual, the transducer will also have a resistive mechanical impedance, and that the mechanical resistance will be inversely proportional to the electrical resistance.

Actual transducers differ considerably from the idealized one described

TRANSDUCERS

loudspeaker (both these values being for a current of 1 amp in 3 ohm coils). With this force available, it is possible to use a very stiff spring across the transducer to raise its output impedance; movements of the order of 1 mm are possible, the resonant frequency then being of the order of 200 c/s.

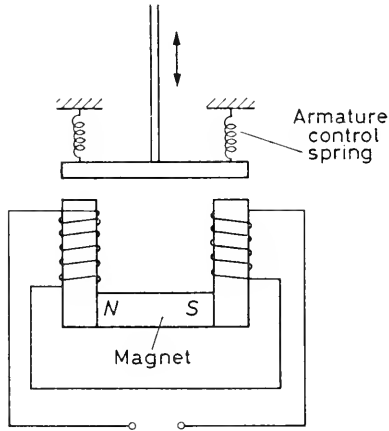


Figure 33.24 Moving-iron transducer with reasonable linearity

Another configuration of this type of transducer is used in the moving-coil meter (*Figure 33.25*). This is capable of only a very small mechanical output, and in its usual form has a very low resonant frequency. An ordinary moving-coil meter has however been adapted as an electrically controlled torsion balance; other similar applications when a small force only is required may suggest themselves.

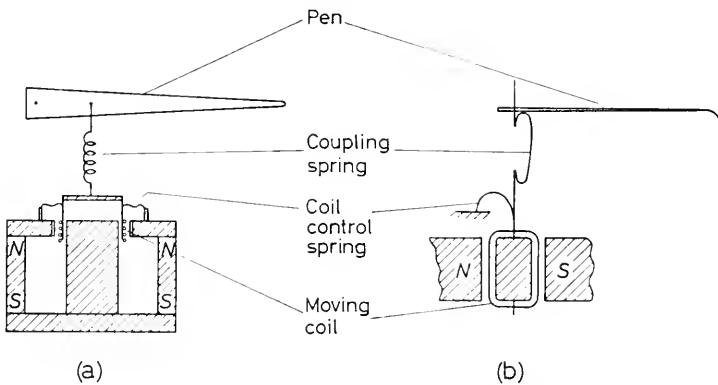


Figure 33.26 Pen recorder: (a) moving-coil loudspeaker type; (b) moving-coil meter type

The moving-coil pen recorder (also known as the penwriter and pen oscillograph) can be made using either the loudspeaker principle or the meter principle; the two configurations are illustrated in *Figure 33.26*. In each case the driving coil is controlled by a spring, and another spring is used to couple the coil to the pen. Viscous damping by means of oil is



Figure 33.25 Moving-coil meter

ELECTROMECHANICAL TRANSDUCERS

sometimes provided while other designs employ the damping effects of eddy currents induced in a metallic coil former. In all cases some additional mechanical resistance is introduced since the driving amplifier has a resistive output impedance.

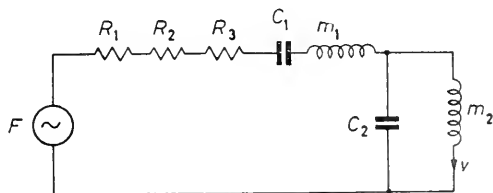


Figure 33.27 Equivalent circuit of pen recorder

The equivalent circuit of the mechanical parts of a pen recorder is shown in *Figure 33.27*. Here R_1 is the mechanical resistance induced by the output resistance of the driving amplifier, R_2 and R_3 are the resistances (if any) due to oil and eddy current damping, m_1 and m_2 are the effective masses of the coil and pen, and C_1 and C_2 are the compliances of the coil control spring and the pen coupling spring. The displacement of the pen is represented by the charge flowing through m_2 . By suitable choice of the components it can be arranged that the sensitivity of the recorder is constant from zero frequency to, in a typical case, 100 c/s.

This adjustment depends, of course, on R_1 having the correct value. It is usual in the design of driving amplifiers for pen recorders to provide a variable degree of negative feedback. The output impedance of the amplifier may thereby be varied (*cf.* Chapter 11), and hence the value of R_1 may be controlled. If R_1 is too small there will be a resonant peak in the response of the recorder and the response to a transient will be oscillatory.

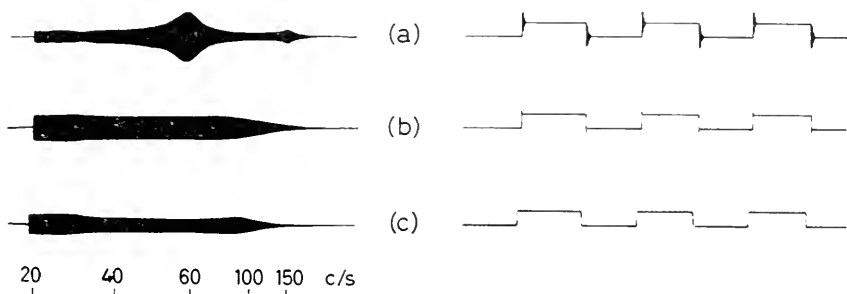


Figure 33.28 Frequency response and transient response of pen recorder: (a) under-damped; (b) correct damping; (c) slightly over-damped

If R_1 is too high the transient response will be over-damped. A technique for adjusting the damping is thus available; a low-frequency square wave is applied to the recorder, and the feedback control adjusted until it is reproduced most faithfully. In *Figure 33.28* the square wave response and frequency response of a pen recorder are shown for three different degrees of damping.

TRANSDUCERS FOR CONTINUOUS ROTATION

All the devices described in this section lie within the field of servomechanism engineering. Since this field is a particularly wide and difficult one, a brief description will be given of only those devices which might have some application in biological research. Thus for a.c. servomotors and generators, rotating amplifiers and special-purpose servo techniques, the reader should consult the extensive literature now available on servomechanisms^{7,8,9}.

Before discussing rotating devices under the headings of m-e and e-m transducers, systems which might formally be called mechano-electro-mechanical will be described under their more usual name of 'synchronous links'.

Synchronous links

It is frequently necessary to transmit shaft rotations over paths so long or tortuous that flexible shafts or belts are impracticable. For these purposes pairs of electrically interconnected devices known as synchronous links are available. These may be used in very low-power applications, such as for driving indicating pointers, or where considerable power is transmitted; the underlying principles are similar.

D.c. synchronous links—The Desynn is illustrated in *Figure 33.29*. The

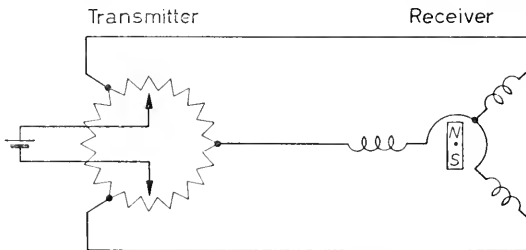


Figure 33.29 The Desynn

variation of current in the three windings of the receiver due to the rotation of the transmitter shaft causes the receiver shaft to turn synchronously. Desynns are useful only as low-power links, and are commonly used for indicating the position of a shaft remotely.

The M-motor (*Figure 33.30*) uses a similar receiver configuration, but moves discontinuously in 12 steps per revolution. Quite high power may be transmitted by devices of this type. Torques of 150 g-cm and speeds up to 200 rev/min can be achieved with an M-motor of 2 in. diameter and 2 in. long. When step-wise motion is of no consequence, as for instance when a considerable step-down gear ratio is used, the M-motor provides a most convenient source of remotely controlled power.

A.c. synchronous links—The principle of the majority of a.c. synchronous links is illustrated in *Figure 33.31*. The currents induced in the stator windings of the transmitter give a magnetic field configuration in the receiver which pulls the two shafts into synchronism.

When used as a low-power link for indication only, a simple and light

TRANSDUCERS FOR CONTINUOUS ROTATION

construction is used in the receiver. However, considerable power may be transmitted if identical units are used at either end of the link; due to the symmetry the power may be transmitted in either direction.

The terminology of a.c. synchronous links is exceedingly confused, due to differences of nomenclature between Great Britain and the U.S.A., between the Armed Services and between different manufacturers. In the

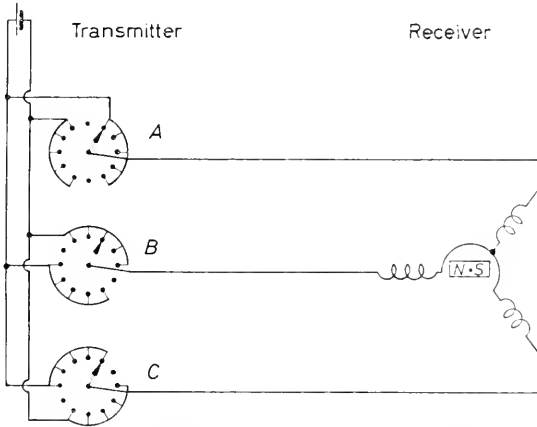


Figure 33.30 The M-Motor. A, B and C are ganged switches or cam-operated contacts

U.S.A. the term 'synchro' covers all sizes of these devices; in addition, nearly every manufacturer has his own trade name (e.g. Selsyn, Autosyn, Diehlsyn, Teletorque, Asynn). In Great Britain it is usual to call low-power units magslips, and large high-power units Selsyns. The details given here will refer primarily to magslips, but will apply approximately to synchros of equivalent size.

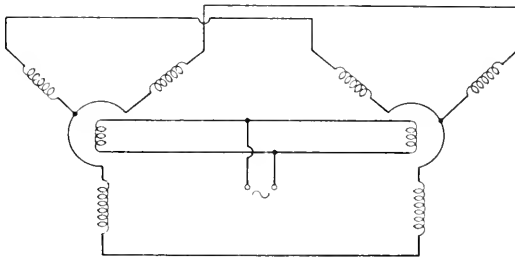


Figure 33.31 A.c. synchronous link—general principle

With an indicating link, using the special receiver designed for this purpose, the accuracy of the position information is about ± 1 degree, but the receiver cannot supply any mechanical output. A large number of receivers may be driven from one transmitter. The accuracy of two similar units used in a synchronous link depends on the mechanical load; on no load ± 1 degree is possible, while errors of up to 20 degrees will occur if the link is operated at its maximum load. Magslips are made in three sizes: $1\frac{1}{2}$, 2 and 3 in. diameter; the corresponding maximum useful torques in a

TRANSDUCERS

synchronous link are approximately 100, 200 and 600 g-cm. Several receivers may be driven from one transmitter, provided that the total torque transmitted is not excessive.

Magslips operating on 50 c/s supplies are commonly wound for 50 V, but units for other frequencies and voltages are available. A very wide range of different windings is available for special purposes, but manufacturers' literature¹⁰ must be consulted for details.

Mechano-electrical rotating transducers

It will by now be apparent that all the transmitters of synchronous links are m-e transducers, transforming information about shaft position into electrical form. It will also be apparent that the electrical information is in a most inconvenient form: that of the ratios between three voltages. For example, although the output of one phase of a magslip may be phase-sensitive rectified to give a voltage proportional to the sine of the angle of rotation, it is unusual to use the electrical output of a single m-e transducer as a guide to the position of a shaft. Such transducers, used in pairs, commonly provide information about the *difference* in position between two shafts, but this will not be discussed here.

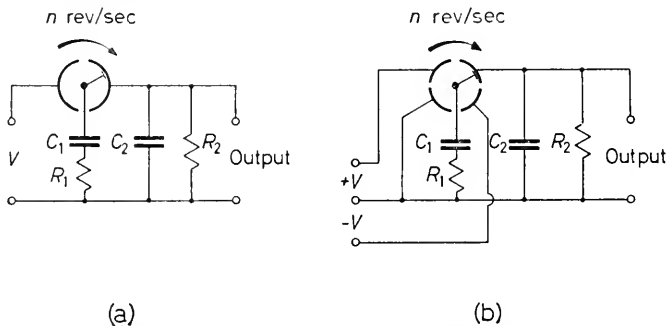


Figure 33.32 Commutated capacitor tachometer: (a) output independent of direction of rotation; (b) output changes sign as rotation reverses

One form of rotating transducer which is in common use is the analogue of the moving-coil m-e transducer described earlier; it produces a voltage output proportional to the speed of rotation of a shaft. Such rotating velocity transducers are called 'tachometers'.

A d.c. generator with a constant field gives an output voltage proportional to its speed, and which reverses in sign as the direction of rotation changes. The constant field may come either from a permanent magnet or from a field winding fed with constant current. Linearity of a high order is possible with a well-designed machine provided that negligible current is drawn from the output.

A small and simple device, capable of good accuracy as a tachometer, is the commutated capacitor tachometer, commonly known as the 'bucket machine'¹¹. The circuits of *Figure 33.32* show: (a) a device whose output does not change sign with direction of rotation; (b) a rotation-sensitive device. The capacitor C_1 is charged once per revolution through the protective resistance R_1 . C_2 is made much larger than C_1 so that in due course

TRANSDUCERS FOR CONTINUOUS ROTATION

very nearly all the charge of C_1 is transferred into it. The charge transferred per second is thus nVC_1 , i.e. the mean current flowing in R_2 is nVC_1 . The mean voltage across R_2 will be nVC_1R_2 ; the output is thus proportional to shaft speed. For accuracy it is necessary that the time-constant C_1R_1 should be much less than (say 1/10 of) the dwell-time on the contacts; this ensures adequate charging. Further, the output voltage must be kept low compared with V , as otherwise all the charge is not transferred to C_2 .

Both generator and capacitor tachometer suffer from ripple on their outputs; some smoothing is often necessary.

Electromechanical rotating transducers

In this section two examples will be considered of the vast number of 'instrument motors' which may be used to convert electrical inputs into rotating outputs.

The simplest arrangement uses a motor with a constant (permanent or wound) field, the armature being driven from the electrical input. Here

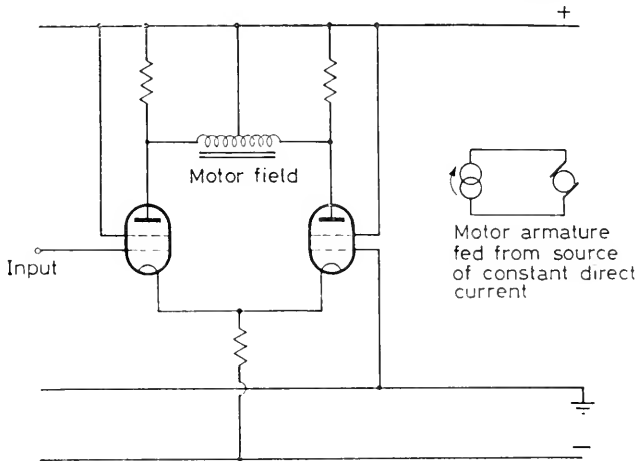


Figure 33.33 Circuit for split-field motor

the speed of the motor will be roughly proportional to the input voltage, and the torque to the input current. Constant-voltage and constant-current feeds thus give approximately constant-speed and constant-torque mechanical outputs, illustrating again the interaction of electrical and mechanical impedances. While large motors may have their armatures fed from batteries of large valves, the split-field arrangement (see below) is more convenient. However for low-power applications it is usually possible to drive the armature of a midget motor (of the type used in toys) directly from a valve or, better, a transistor. One motor which has been successfully used in such an arrangement requires 12 V 0.4 amps to produce its full output of 1/1,000 H.P., and will run at speeds up to 9,000 rev/min. A large output valve or a power transistor will provide this power easily.

For larger motors it is usual to use the electrical signal only to control the motor power, not to provide it, by varying the excitation of a wound field. This field is commonly balanced, giving rise to the term 'split-field'

TRANSDUCERS

motor. A typical circuit arrangement is shown in *Figure 33.33*; a mechanical output of 60 W is controlled by an electrical input of 10 W. The motor output torque is roughly proportional to the field current. Such motors are often structurally combined with tachometer generators; *Table 3* gives an example of the characteristics of such a device.

TABLE 3

<i>Motor-generator type 88</i>		
Motor	{	Field current 25 mA Armature current 0.8 amp Torque 180 g-cm Speed 9,000 rev/min
Generator	{	Field current 30 mA Output voltage 0.01 V per rev/min

CIRCUITS FOR USE WITH STATIC TRANSDUCERS

As has already been noted, static transducers need an external supply of power for their operation, and special circuits to present their output in a convenient form. There are certain fundamental points common to many of the circuits which will be discussed first.

It is desirable that the transducer and its circuit should where possible be symmetrical and balanced. This balance should appear not only in the circuit diagram but also in the mechanical construction of the transducer. Such balanced arrangements are inevitably less subject to the effect of outside disturbances, less liable to drift, and often more linear than their unbalanced counterparts.

The second point concerns only transducers fed from a.c. It is essential for faithful reproduction that the frequency of the energizing supply be considerably higher than any frequency which it is required to study: otherwise, when the output of the transducer is rectified and smoothed to feed, say, an oscillograph, some information of interest will also be removed by the smoothing process. While in theory an energizing frequency of twice the highest signal frequency is all that is necessary, the smoothing filters needed are so complex that the use of this theoretical frequency is not practicable. At least ten times the maximum signal frequency is recommended to energize a.c. transducers.

Circuits for variable resistance transducers

All circuits which are used with variable resistance transducers are variations of the Wheatstone bridge circuit shown in *Figure 33.34*. Sometimes all the components of the bridge are not obvious; R_1 may be the transducer, R_2 a resistor, while R_3 and R_4 may be part of the shift circuits of the measuring oscillograph. It is desirable in order to achieve balance that R_1 and R_2 should be of the same order of magnitude, and of similar construction. It is often possible to make the transducer provide both R_1 and R_2 , as in the liquid or wirewound potentiometer. When R_1 is, for

CIRCUITS FOR USE WITH STATIC TRANSDUCERS

example, a strain gauge, R_2 may be a similar 'dummy' gauge, connected with similar wires and situated in the same environment, but not subjected to strain.

The bridge may be fed with either d.c. or a.c., except when using a liquid potentiometer, when a.c. is usually used to avoid polarization of the electrodes. No difficulty arises when using d.c., as when the bridge is unbalanced one way or the other the output signal is positive or negative. With a.c. energization the output signals on the two sides of balance are 180 degrees out of

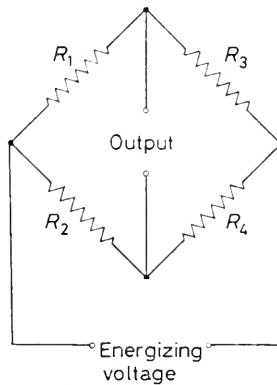


Figure 33.34 Wheatstone bridge circuit

phase with each other. With simple rectification this phase distinction disappears, leaving the direction of unbalance ambiguous. This may be resolved in one of two ways: (a) the bridge may be deliberately unbalanced, so that the output never goes through zero; (b) a phase-sensitive rectifier (Chapter 6) may be used. While (a) is very simple, (b) confers advantages in simplicity of smoothing circuits and in overall signal-to-noise ratio.

Circuits for variable reactance transducers

It is again possible to use suitable bridge circuits¹² with variable capacitance or variable inductance transducers; all the considerations of the previous section about balanced design, use of phase-sensitive rectifiers, etc. apply. It is essential, of course, to use a.c. energization. The differential transformer is an example of a concealed bridge circuit; the bridge elements are the reluctances of the elements of the magnetic circuit, the output being derived from the flux in the 'diagonal' member.

Variable reactance transducers can conveniently be used in an entirely different type of circuit. The variable element forms part of a tuned circuit, whose resonant frequency is thus varied by the mechanical input. This frequency is then determined continuously and presented in a suitable form as the output of the transducer system. The variable tuned circuit may conveniently be used as the controlling element of an oscillator, the changes of whose frequency are converted into proportional voltages by a discriminator circuit of the type used in frequency-modulation radio receivers. Two examples will be given; many other circuits will occur to the reader who studies f.m. receiver design practice^{13,14}.

TRANSDUCERS

The tuning inductor of the oscillator in *Figure 33.35* has an open magnetic circuit, and is varied by the movement of a mumetal armature. The output of the oscillator is converted into square waves by the limiting valve, and

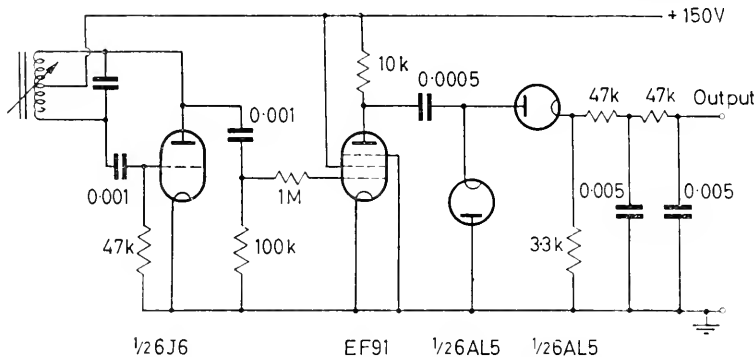


Figure 33.35 Circuit for variable inductance transducer

these square waves feed a 'pulse counter discriminator'¹⁵ which gives a d.c. output voltage proportional to the frequency of the oscillations. Once again the frequency of oscillation must be high compared with the signal frequency.

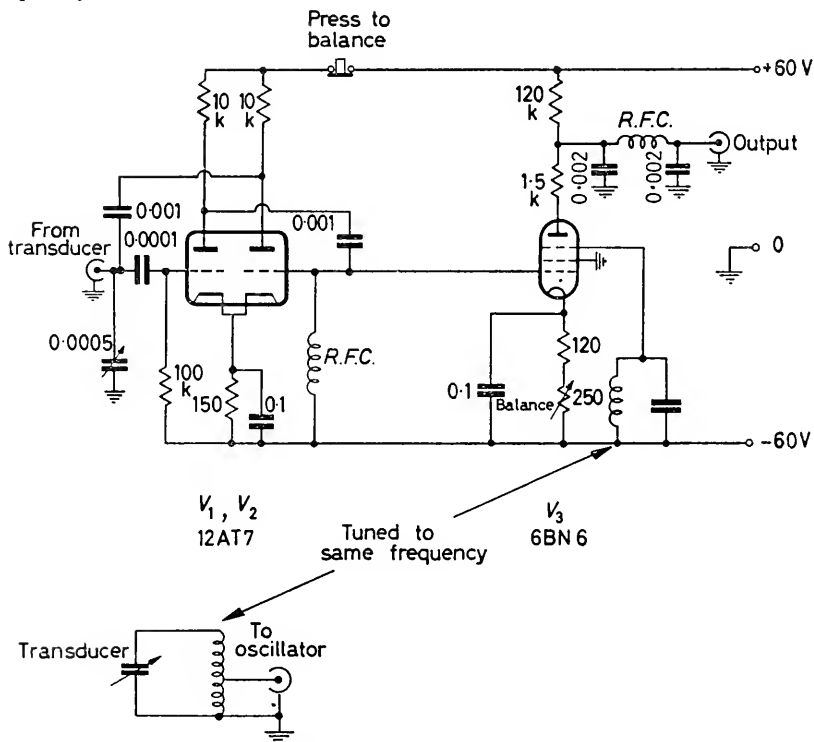


Figure 33.36 Circuit for use with capacitance transducer

TRANSDUCERS WITH FEEDBACK

A circuit suitable for use with the capacitance transducer manometer (page 505) is given in *Figure 33.36*. The manometer capacitance determines the frequency of the oscillator V_1 and V_2 . This frequency is converted into a voltage output by the 'gated beam discriminator' V_3 ¹⁶. An output of 50 mV per cm of water pressure (representing 10 V per pF capacitance change) is obtainable, with very good stability. The operating frequency is about 1 Mc/s; such high frequencies are necessary with transducers of low capacitance or inductance, as in order to achieve high sensitivity the transducer should form the major part of the tuning reactance.

The exact mode of operation of the discriminators described is beyond the scope of the present work; the references quoted should be consulted for further details.

TRANSDUCERS WITH FEEDBACK

The principles of negative feedback have been described in Chapter 11 where it was shown that its use can bring the following advantages: (a) improved linearity; (b) improved frequency response; (c) change of input or output impedance.

Negative feedback may similarly be applied to transducers with equivalent advantages. It is not possible to cover all applications of negative feedback to transducers, but examples will be given which illustrate the principles. As in all feedback circuits, care is needed to avoid continuous oscillation due to a large number of phase shifts in the loop. This effect and its cure are discussed on pages 165 and 166.

A self-balancing mechano-electrical transducer for force measurement

It has been noted that an m-e transducer measuring applied force must have as high a mechanical input impedance as possible. This can be

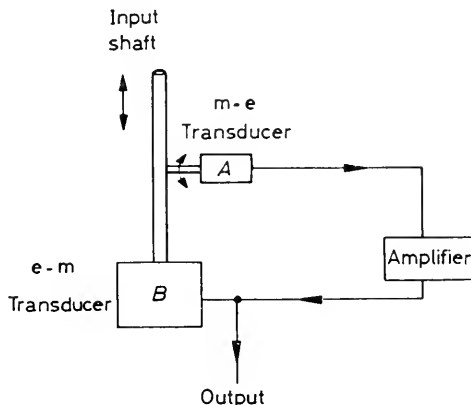


Figure 33.37 Feedback force-transducer—block diagram

achieved with any transducer if a stiff spring is placed across it: the sensitivity will then be low, and considerable amplification will be needed. In the device to be described this amplification is used to improve the properties of the transducer by the application of negative feedback.

TRANSDUCERS

In *Figure 33.37* the applied force is imposed on two transducers mechanically interconnected. Transducer *A* is an m-e transducer which detects any movement of the assembly from its rest position. The output from *A* is amplified and fed to the e-m transducer *B* so as to annul any movement. This arrangement will be recognized as a feedback system. From the type of analysis presented in Chapter 11 it follows that: (a) the transducer has a very high mechanical input impedance, i.e. it does not 'give' appreciably under the applied force; (b) the applied force is very nearly equal to the force produced by the e-m transducer, and can therefore be measured from the electrical input to this transducer; (c) the characteristics of the m-e transducer do not affect the measurement.

Since highly linear e-m transducers are easily constructed, and m-e transducers of high sensitivity can be made if linearity is not important, the combination described has many practical advantages.

An electromechanical transducer for controlled displacements

It is difficult to design an e-m transducer whose output displacement is accurately proportional to its electrical input if the range of displacement must be large. It is even more difficult under these conditions to make the movement independent of load (i.e. high output impedance).

By an arrangement (*Figure 33.38*) similar to that of the previous section

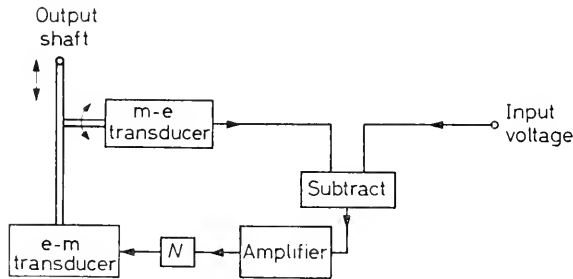


Figure 33.38 Feedback system for controlled displacements

it is possible to obtain a displacement proportional to an input voltage, and virtually independent of the load. Now the m-e transducer must be adequately linear, while the characteristics of the e-m one are more or less unimportant. A suitable phase-advance stabilizing network *N* will however be essential to avoid oscillation of the system (see page 165); its exact design will depend on the mechanical properties of the e-m transducer.

The circuit will so adjust itself that the difference between the input voltage and the output of the m-e transducer is very small; the displacement of the whole system will thus reproduce accurately the waveform of the input voltage.

The velodyne

This circuit¹⁷ is intended to give an output shaft-rotation, whose angular velocity is proportional to an input voltage. It has two principal uses: (a) to give a rotating motion where speed may be readily controlled (manually or automatically)—this speed being virtually independent of load (i.e. high

THE USE OF TRANSDUCERS IN BIOLOGY

output impedance); (b) to carry out integration of a voltage on a long-time scale. Since the angular velocity of the shaft is proportional to the input voltage, its number of revolutions (determined with a mechanical counter) measures the time integral of the voltage. Thus, for example, a linear photocell connected to a velodyne could give a measure of the total light energy falling on the cell over a period of as long as a year—a measurement otherwise very difficult to make.

It will be seen from *Figure 33.39* that motor and tachometer generator

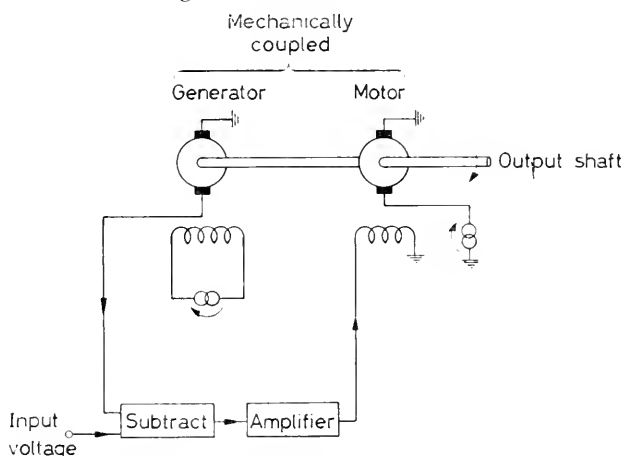


Figure 33.39 The velodyne—block diagram

are mechanically coupled, while the difference between the input voltage and the tachometer output is amplified to drive the motor. Since this difference is therefore held small, the angular velocity of the shaft must be closely proportional to the input voltage.

THE USE OF TRANSDUCERS IN BIOLOGY

In this section some examples will be given illustrating the use of transducers in biological research.

A capacitance transducer manometer

An application of the capacitance m-e transducer which is notable both for the elegance of its construction and for the completeness of the mathematical and experimental performance analysis is due to Hansen¹⁸⁻²⁰. His design for a capacitance manometer is shown diagrammatically in *Figure 33.40*. A change of capacitance of about 0.05 pF is produced by a pressure change of 1 cm of mercury. Using a circuit similar to that described on page 503, this pressure change can be arranged to give an output of 0.5 V.

The manometer with hypodermic needle can have a resonant frequency as high as 100 c/s, with nearly critical damping. This response is more than adequate for delineating accurately the time-course of human blood pressure (the purpose for which the instrument was designed), and is

TRANSDUCERS

likely to be sufficient for measuring pressure variations in nearly all biological systems.

In addition to describing the manometer, Hansen²⁰ also reviews previous designs, most of which are m-e transducers. The same reference contains

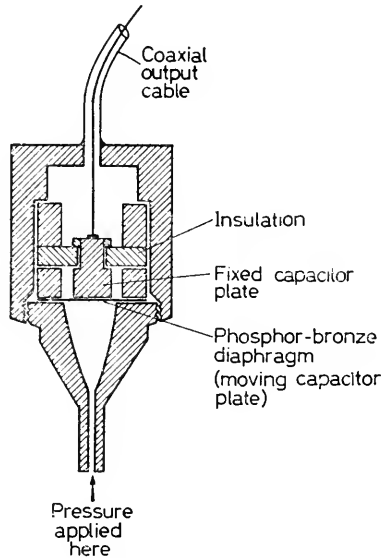


Figure 33.40 Hansen capacitance manometer

some very sophisticated examples of the use of electrical equivalent circuits in the solution of mechanical problems.

Investigation of the vibration sensitivity of spiders

In an attempt to discover how a spider detects the presence of its prey on the web, Dr. D. A. Parry has carried out some preliminary experiments on vibrations of the web, using transducers designed and constructed by J. A. Popple*. These experiments illustrate well some of the difficulties encountered in the application of transducers to biological material; they also illustrate how a simple transducer, designed specifically for the one purpose, may lead to results otherwise obtainable only with delicate and complex apparatus.

The first experiments sought to measure the vibrations induced in a web by a fly caught on it. A very small and light differential transformer of the type shown in *Figure 33.17* was used. The armature, which was brought into contact with the web, was a mumetal wire about 0.1 mm diameter and 5 mm long, flexibly mounted through a sheet of rubber 0.05 mm thick. For easy recording of vibration frequencies up to 1,000 c/s an energizing frequency of 10 kc/s was used, with the output of the transducer feeding the Y plates of an oscillograph. A moving-film camera gave a record in

* I am indebted to Dr. Parry and Mr. Popple for allowing me to describe this preliminary (and unpublished) work, particularly as both successful and unsuccessful applications of transducers are mentioned.

the form of an amplitude-modulated strip, whose envelope described the vibrations experienced at the transducer.

While some results were obtained by this method, it was found that the transducer was much too stiff (i.e. was of too high mechanical impedance) and as a result the web near the point of contact was soon torn. However delicately the transducer armature was constructed, the impedance it presented at the web was always so much greater than the impedance of the web fibres that the resulting stresses damaged the web. Here, then, is a case where the biological material is so delicate and the forces involved so small that not even the most refined m-e transducer is likely to be of value.

A similar result was obtained when an attempt was made to induce vibrations in the web with a moving-coil e-m transducer (a moving-coil loudspeaker with the cone removed, fitted with a stylus). The relatively high output impedance of the transducer again caused damage to the web. However, a moving-iron transducer of very low mechanical impedance, and of great simplicity, finally proved successful. A very small iron turning was placed on the web, and subjected to an alternating magnetic field from an iron-cored solenoid driven from an oscillator. This placed virtually no constraint on the web and made it possible to originate controlled vibrations at any point in it.

A variable mechanical load

It is known that the flight muscles of certain insects, when stimulated, can display cyclical changes of length at a much higher frequency than that of the stimulus. The frequency and amplitude of these oscillations is dependent on the external mechanical load applied to the muscle; this load in the living insect is of course the wing, and the oscillation produces the wing beat.

The first experiment demonstrating this effect in an isolated muscle preparation was due to Boettiger; he attached the muscle to a lever carrying a mass to provide inertia, and had a friction arrangement to give damping. To gather more data in the relatively short life-time of the preparation, the author has devised an apparatus²¹ whereby mechanical loads can be generated electronically. Any combination of mass, viscosity and compliance can be produced, the amount of each being instantaneously variable by potentiometers in the electronic circuits. In addition any of the components may be made negative, so that mass, compliance or viscosity present in the preparation can be cancelled out.

Figure 33.41 shows the block diagram of the equipment. The muscle preparation is attached to the output shaft of a moving-coil e-m transducer. A vane is attached to the same shaft and moves in the path of a beam of light to a photocell. The output voltage of the photocell is thus made proportional to the displacement x of the output member. This voltage is then differentiated twice, giving voltages proportional to \dot{x} and \ddot{x} ; voltages proportional to $-x$, $-\dot{x}$ and $-\ddot{x}$ are generated by inverting amplifiers. With the arrangement shown, fractions α , β , γ (which may be positive or negative) of the three voltages representing x , \dot{x} and \ddot{x} are selected and

TRANSDUCERS

added together. This sum is converted into a current in the moving coil, and hence to the force F on the output member.

We therefore have that

$$F \propto \alpha x + \beta \dot{x} + \gamma \ddot{x}$$

By comparison with the equation

$$F = m\ddot{x} + R\dot{x} + \frac{1}{C}x$$

it follows that the parameters α , β , γ , controlled by the potentiometers, set the values of the apparent compliance, viscous resistance and mass

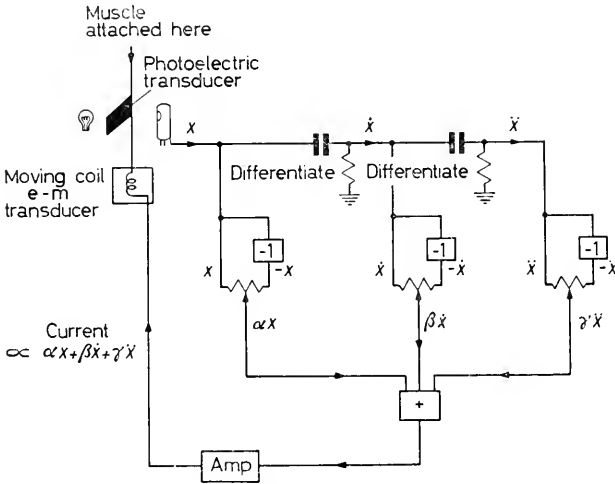


Figure 33.41 A variable mechanical load—block diagram

exhibited by the output member. It should be noted that the mass thus generated has the usual inertial properties, but is not subject to gravitational forces. Any static force required can be fed into the adding circuit in the form of a constant voltage.

This circuit shows once more how the output impedance of a transducer is modified by feedback (in this case either positive or negative). This application of transducers is unusual for the following reason. The complex arrangement of two transducers and associated electronic circuits is not employed to 'convert quantities of one type into equivalent quantities of another type' (page 471); it merely serves to modify the mechanical parameters of the input shaft.

CONCLUSION

The biologist who decides to use a transducer should first ask himself whether the experiment could not be carried out better with a lever and a piece of string. Only when he is satisfied that a transducer is necessary, either for reasons of sensitivity or of convenience, should he proceed to choose a suitable design. The impedance of the transducer must be considered at this stage. If displacement is to be measured, how much loading will the preparation stand? If an electromechanical transducer is to be

REFERENCES

used to apply force to a preparation, does it matter if the force changes as the preparation moves? Impedance considerations will narrow the field of choice considerably.

The frequency response of the transducer must now be matched to the experiment. Must the transducer respond to steady deflections? What is the highest frequency present in the signals under study? Whether the transducer is to be static or dynamic, and how high its resonant frequency must be, can now be decided.

The design of the transducer will now be fairly clear, and it may be either obtained or constructed. It is then essential that it should be tested *under the experimental conditions*. Generally it is simpler to determine the transient response than the frequency response; a sudden electrical signal can be applied to an electromechanical transducer, while a suitable mechanical shock to a mechano-electrical transducer can usually be arranged. An estimate of the resonant frequency and damping follows from experiments of this type. It is now possible to confirm that the properties of the transducer will not unduly modify or distort the signals.

Tests of sensitivity and freedom from external influences are straightforward. Electrical or magnetic screening and possibly isolation from mechanical vibrations may be necessary.

Finally it may fairly be said that if the properties of the transducer and the extent to which it distorts the experimental results are not understood, it would be better to do the experiment another way.

REFERENCES

- ¹ CANAWAY, R. J. *et al.* *Electron. Engng.* 27 (1955) 103
- ² FARRAR, J. T., ZWORYKIN, V. K. and BAWN, J. *Science* 126 (1957) 975
- ³ KING, R. W. P., MIMNO, H. R. and WING, A. H. *Transmission Lines, Antennas and Wave Guides* McGraw-Hill. 1945
- ⁴ TERMAN, F. E. *Radio Engineering*, p. 75. McGraw-Hill. 1951
- ⁵ Radio Corporation of America. Data Sheet *Type 5734 Mechano-electronic Transducer*
- ⁶ PASCOE, J. E. *J. Physiol.* 128 (1955) 26P
- ⁷ LAUER, H., LESNICK, R. and MATSON, L. E. *Servomechanism Fundamentals* McGraw-Hill. 1947
- ⁸ WEST, J. C. *Servomechanisms* E.U.P. 1953
- ⁹ PORTER, A. *Introduction to Servomechanisms* Methuen. 1950
- ¹⁰ Muirhead Ltd. Catalogue
- ¹¹ WILLIAMS, F. C. and UTTLEY, A. M. *J. Instn. elec. Engrs.* 93 Pt. IIIa (1946) 1268
- ¹² OWEN, D. *Alternating Current Measurements* Methuen. 1937
- ¹³ JOHNSTONE, G. G. *Wireless World* 63 (1957) 8, 70, 124, 275, 378
- ¹⁴ STURLEY, K. R. *Radio Receiver Design*, p. 294. Chapman & Hall. 1947
- ¹⁵ SCROGGIE, M. G. *Wireless World* 62 (1956) 158
- ¹⁶ JOHNSON, L. W. *Wireless World* 63 (1957) 23
- ¹⁷ WILLIAMS, F. C. and UTTLEY, A. M. *J. Instn. elec. Engrs.* 93 Pt. IIIa (1946) 1256-1274
- ¹⁸ HANSEN, A. T. *Acta physiol. scand.* 19 (1950) 306
- ¹⁹ HANSEN, A. T. *Acta physiol. scand.* 19 (1950) 333
- ²⁰ HANSEN, A. T. *Pressure Measurements in the Human Organism* Copenhagen Teknisk Forlag. 1949
- ²¹ MACHIN, K. E. To be published.

RELAYS AND RELATED MECHANISMS

K. E. MACHIN

INTRODUCTION

A relay is a device in which an electrical input produces a mechanical movement, this movement opening or closing electrical connections. It is thus an electro-mechano-electrical transducer with only two positions—unoperated and operated. The mechanical movement is usually produced by an electromagnet which moves a number of contacts to set up the controlled electrical connections.

The two main uses of relays are: (a) controlling a large amount of electrical power with a small amount, or from a remote position; (b) setting up and controlling complicated circuit interconnections. Numerous examples of (a) are described in Chapter 29. An example of (b) is the use of relays in a large power supply unit to do the following operations: (1) 30 seconds after the unit is switched on, switch on the negative HT; (2) as soon as the negative HT is on, switch on the positive HT; (3) if at any stage the negative HT fails, switch off positive HT and ring an alarm bell continuously; and (4) if the positive HT fails, ring the alarm bell intermittently.

As a further example, a relay circuit can be used to record on a smoked drum only every fourth impulse received from the closing of a pair of contacts.

In this Chapter the fundamentals of relay design are not discussed since the reader is never likely to construct his own relays. Instead, practical information is given about the use of relays and about their properties and limitations. Since many thousands of different designs of relay exist it is quite impracticable to discuss them all. Detailed descriptions are given of those relays most likely to be useful, while a few general characteristics of the others are mentioned.

Attention has been concentrated largely on the Post Office 3000 type relay, since this is readily available, standardized and of proved reliability and usefulness. The general features of relay design will be illustrated with reference to it, and much of the information given will be relevant to relays of other design.

NOTATION AND SYMBOLS

Relay circuits were formerly drawn with the coils and contacts of each relay near together. A circuit drawn in this way is shown in *Figure 34.1*. Even a simple circuit of this type looks confused; more complicated circuits drawn using the 'attached-contact notation' are almost unreadable.

NOTATIONS AND SYMBOLS

It therefore became usual to draw the coils and contacts of a relay in the most convenient place on the diagram, and to associate them only by means of a code of symbols. The circuit of *Figure 34.1* is thus transformed into that of *Figure 34.2*. The number within the rectangle represents the resistance

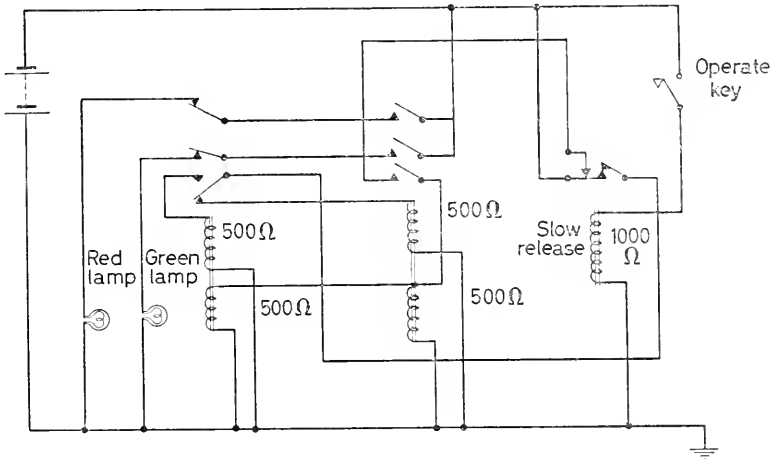


Figure 34.1 Relay circuit drawn with attached-contact notation

of the coil in ohms, while the code $A/3$, for example, identifies the relay as A and indicates that it has 3 contacts. Special properties (e.g. slow release or high speed) of the coil are shown symbolically. The contacts are labelled

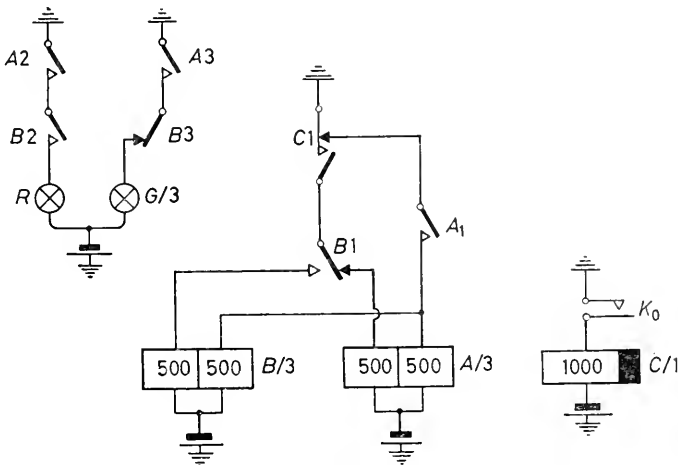


Figure 34.2 *Figure 34.1* circuit redrawn with detached-contact notation

with the relay's code letter, and their serial number (e.g. A_2 , the second contact set on relay A).

The usefulness of the 'detached-contact notation' cannot be emphasized too strongly. It provides a complete and convenient symbolism, and once it has become familiar relay circuits are easy to follow. The perpetration

RELAYS AND RELATED MECHANISMS

of diagrams similar to *Figure 34.1* is then both unnecessary and unfor-
givable.

The symbols used in relay circuits are summarized in *Figure 34.3*. Contacts

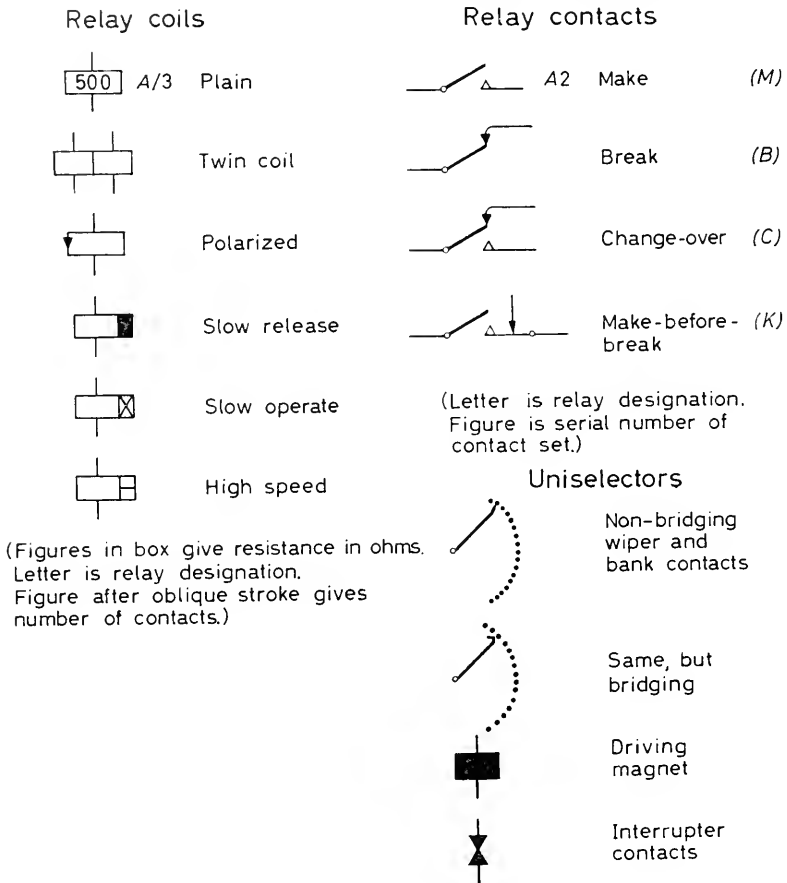


Figure 34.3 Symbols used in relay circuits

are always drawn in the unoperated position; thus a 'make' contact is drawn open, and a 'break' contact closed. In a complicated circuit the symbols for 'earth' and 'earthed battery' are used wherever a connection to one or other pole of the supply occurs (see *Figure 34.2*); it is of course not thereby implied that as many batteries as symbols are provided.

THE POST OFFICE 3000 TYPE RELAY

General construction

The components of a 3000 type relay are illustrated in *Figure 34.4*. A soft iron yoke surrounds a detachable coil wound on an iron core. An L shaped armature rests on a knife-edge machined on the yoke, and is retained by

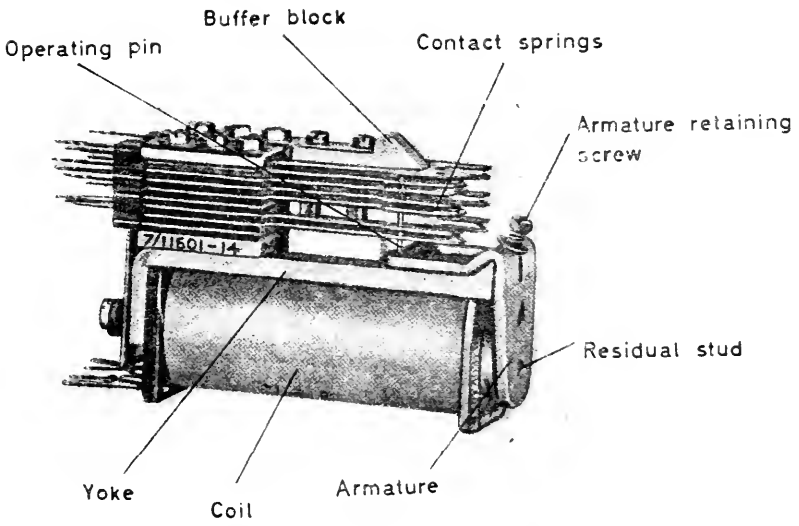


Figure 34.4 The P.O. 3000 type relay

(By courtesy of J. Atkinson and Sir Isaac Pitman and Sons, Ltd., from *Telephony*, 1948)



a small screw and spring. A residual stud or screw prevents the armature sticking to the coil face. The contact springs are mounted above the armature, and are operated by it through projecting pins: lugs on some of the springs engage on a moulded buffer block. The relay is best mounted on its side so that accumulated dust may fall clear of the contacts.

The coil unit

The coil assembly is held in place by a single nut and may be removed even when the relay is mounted. The windings terminate in tags on the rear bakelite cheek; up to five tags may be provided. When the coil is viewed from the rear with the tags at the bottom, reading from left to right the tag positions are designated *a, b, c, d, e*. Details of the windings printed on the coil are correspondingly lettered. Tags connected to the inner end of a winding are usually coloured red; it is customary to connect them to the earthy end of the supply.

Single and twin coil relays are common; triple and quadruple coils are available for special purposes. Solid copper 'slugs' of three sizes— $\frac{1}{2}$, 1 and $1\frac{1}{2}$ in.—may be provided at either the armature end or the heel end of the coil; the use of slugged relays is discussed in a later section.

The maximum power dissipation of a relay coil is 6 W. The insulation of a coil is rated to withstand 50 V to earth, but this rating is so conservative that 250 V may be applied with little likelihood of failure. Furthermore, the relay is usually insulated from its mounting plate. It is, of course, good practice to arrange the coil in the earthy end of its operating circuit.

Contact springs and buffer block

The contact springs of a 3000 type relay are nickel silver and usually 0.014 in. thick. For special applications 0.012 in. springs can be obtained. Four types of contact are available, as shown in *Figure 34.5*. Contacts are assembled into groups, each of which can contain a maximum of nine springs. Thus each group can provide up to four makes or breaks or three change-overs. Two such groups are mounted on each relay; these groups are 'handed' so that a left-hand assembly cannot be used on the right-hand side of a relay.

Various contact materials are available. For currents up to 300 mA silver is used, and twin contacts (*Figure 34.6*) to reduce open-circuit failures are standard. Platinum contacts of similar construction can be used up to 1 amp; such contacts are identified by a notch in the tip of the spring. The rated voltage for contacts of these types is 50 V, but no trouble is likely to be encountered in operating them up to 250 V. For heavy currents silver-nickel (or, rarely, tungsten) cylindrical contacts are used. These are rated at 250 V, 8 amps.

All the foregoing contact assemblies are insulated with bakelite spacers. For very high voltages pillar-type insulators are available, with contacts rated at 1 kV 2.5 amps or 2 kV 1.25 amps.

When contacts with low contact pressure are used on very low voltages, for example at the input of an amplifier, it is found that the oxide film on the contacts does not always break down: the resistance between the contacts will then be high. To avoid this effect it is sometimes possible to arrange

RELAYS AND RELATED MECHANISMS

for a d.c. voltage of at least 50 V to be applied to the contacts through a high resistance ($\sim 50 \text{ k}\Omega$) in addition to the small signal voltage.

Projecting from the side of each non-moving (or 'buffer') spring is a small tag which engages with a step on the buffer block. This serves two purposes: (a) by pre-tensioning the buffer spring on to the buffer block it is possible

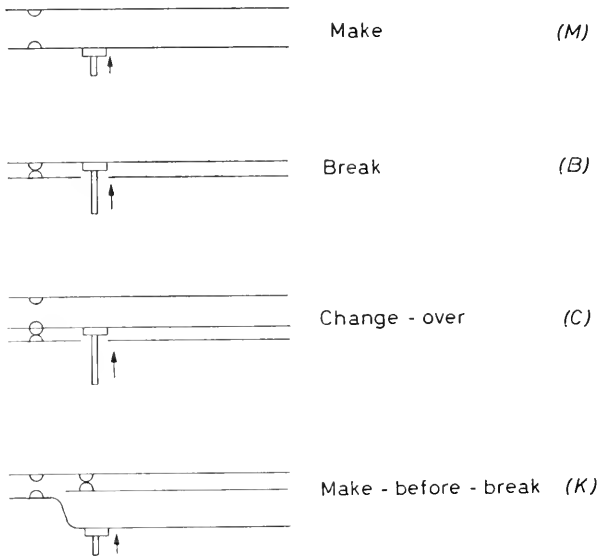


Figure 34.5 Types of relay contact

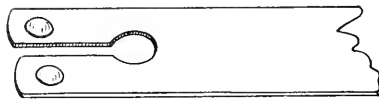


Figure 34.6 Spring with twin contacts

to ensure adequate contact pressure; (b) if the spring is kept straight the position of the buffer block steps determine the contact clearance. It is largely from the buffer block that a relay achieves reliability; operation with unsupported springs can only lead to trouble.

Buffer blocks are available with 1, 2, 3 and 4 steps on each side, and with 3 steps on one side and 4 on the other. All available contact assemblies can thereby be buffered.

Adjustment

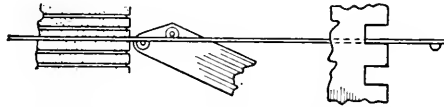
For reliable operation a relay must be properly adjusted. For certain applications it may be necessary to use light spring tensions and short armature movements; it must, however, be understood that reduced power handling capacity of the contacts and greater probability of failure due to dust will be the inevitable results of unorthodox adjustments of this type.

THE POST OFFICE 3000 TYPE RELAY

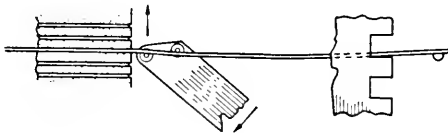
A relay set up to within ± 10 per cent of the figures given below may have a life between adjustments of a million operations.

Before attempting to adjust a relay, the coil retaining nut should be checked for tightness. The armature travel should then be adjusted to 0.030 in. by bending the armature suitably. A special tool is available for this purpose, but a light hammer, delicately used, is an acceptable substitute.

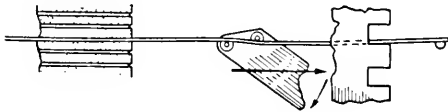
For contact adjustment two special tools are needed: (a) A spring balance to measure the force at the contact tip; if a proper relay balance is not available it is possible to use a piece of leaf spring suitably mounted with a



Position of special spring adjuster



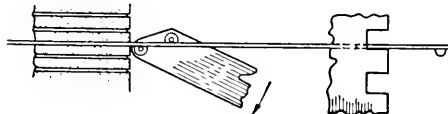
Bending the spring and pressing upwards



Stroking the spring



Bow in spring



Final set in spring

Figure 34.7 Tensioning a contact spring

(By courtesy of J. Atkinson and Sir Isaac Pitman and Sons, Ltd. from Telephony, 1948)

scale which converts deflection into force at the tip. (b) A two-pin spring stroking tool; while it is possible to adjust a relay with a pair of flat-nosed pliers, the correct tool makes the process much more straightforward. The buffer springs should be adjusted first: they should require a force of 20 g wt to lift them clear of the buffer block. They should also be straight so that the

RELAYS AND RELATED MECHANISMS

contact gaps are correct. The process of increasing the tension in a spring is illustrated in *Figure 34.7*; the inverse process is used to decrease tension. When adjusting break buffer springs, the armature should be operated manually to lift the lever spring clear.

The lever springs are next adjusted: make lever springs should bear on the armature or on the contact set below so that a force of 6 g wt is needed to lift them. Break lever springs should be tensioned until they lift the adjacent buffer spring clear of the buffer block when the relay is unoperated. Both tips of twin contacts must touch simultaneously, and cylindrical heavy-duty contacts must meet squarely. Necessary adjustments can be made with the two-pin tool.

As a final test, the relay armature should be operated by hand and checked for smoothness of movement. The contacts should be inspected at the same time; in the unoperated condition all lower buffer springs should be clear of the block, while in the operated position the upper buffer springs should be lifted clear. All springs should be straight, and all contact clearances adequate; furthermore, all break contacts should open before any make contact closes.

Operating power

It is necessary when designing relay circuits to know what power is required for a relay: (a) to operate; (b) not to operate; (c) to hold, once operated;

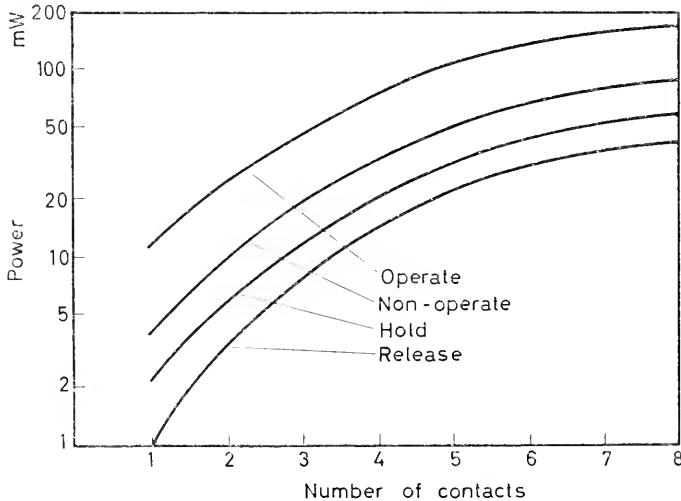


Figure 34.8 Power requirements of 3000 type relay

and (d) to release. The power appropriate to each of these conditions depends on the coil configuration, the number and type of contacts, the thickness of the contact springs and the size of the residual gap. Estimating the power in any particular case is a straightforward though laborious process, for further details of which Atkinson¹ should be consulted. Since, however, for satisfactory operation of relays large factors of safety are usually applied, it is possible to quote results which, while only correct to

THE POST OFFICE 3000 TYPE RELAY

± 30 per cent, will be sufficiently accurate for all practical purposes. If a relay with special adjustments is to be used, or if marginal operation is required, it is essential to test the sensitivity of the relay experimentally.

For normally adjusted, fully wound 3000 type relays, *Figure 34.8* shows the power required for the four cases described above. For thoroughly reliable operation the power read from the graph should be multiplied by the following factors of safety:

Operate	5	Hold	3
Non-operate	0.2	Release	0.15

For coils which are not fully wound the power must be multiplied by $1/f$, where f is the fraction of the coil occupied by the winding. Strictly, this rule applies only when the winding occupies a length-wise fraction f of the available space; for superimposed windings the inner one has the better power sensitivity, since its resistance per turn is lower than the winding of larger average diameter. Accordingly, for non-fully wound coils the power read from *Figure 34.8* should be multiplied by the following factors:

Twin coil with superimposed equal windings:	
Inner coil	1.1
Outer coil	3.5
Twin coil with balanced windings:	
Either coil	2.0
Single coil with $1\frac{1}{2}$ in. slug	2.7
1 in. slug	1.7
$\frac{1}{2}$ in. slug	1.3

Operate and release lags

The time taken by a relay to operate or release depends on a large number of factors which are best summarized in tabular form (*Table 1*).

TABLE 1

<i>For fast operation</i>	<i>For slow operation</i>	<i>For fast release</i>	<i>For slow release</i>
Relay fed from high-impedance source	Relay fed from low-impedance source	No parallel current paths across relay coil	Relay coil shunted with low resistance
Current many times greater than that required to operate	Only just enough current to operate	Only just enough current	Large current
Light contact spring load	Heavy contact spring load	Heavy contact spring load	Light contact spring load
—	—	Large residual gap	Small residual gap
Coil with bakelite front cheek	Armature-end slug	Coil with bakelite front cheek	Either armature or heel-end slug

For a 3000 type relay with a full contact load, energized from a low-impedance supply with a power sufficient to ensure an operating factor of safety of 5, the operate and release lags will be between 10 and 30 ms. Operate lags of less than 5 ms can be achieved if the relay is fed with a large current from a high-impedance source, while release lags of the same order of magnitude can be obtained with low operating power and large residual gap.

Copper slugs fitted to the coil of a relay slow down its action. If the slug is fitted at the armature end the relay is both slow to operate and slow to release. If, however, a heel-end slug is used, only the release of the relay is delayed. Slugs of three sizes— $\frac{1}{2}$, 1 and $1\frac{1}{2}$ in.—are available to give various degrees of delay. With relays in standard adjustment operate lags of about 100 ms and release lags of up to 500 ms can be obtained with $1\frac{1}{2}$ in. slugs.

In the section on Relay Circuits it will be shown that the lags of relays can be modified by external circuit arrangements; some methods of achieving long delays are also given.

Marginal adjustment of relays

The current at which a relay will operate is usually rather ill-defined for the following reasons. The spring load increases as the springs deflect, and the tractive force on the armature increases as it approaches the pole piece. If the spring force increases more rapidly than the tractive force, the relay can be stable in a half-operated condition for certain values of the operating current. Whether or not the contacts are closed by this current depends critically upon the contact clearances. However, if the tractive force always increases more rapidly than the spring load, a current sufficient to overcome the resting spring load will always operate the relay fully.

For this reason, when a relay is required to operate on a well-defined current the spring force must be arranged to increase only slightly with deflection. This can best be done by heavily pre-tensioning the lever springs on to the armature, so that the increase of force due to deflection is a small proportion of the total force.

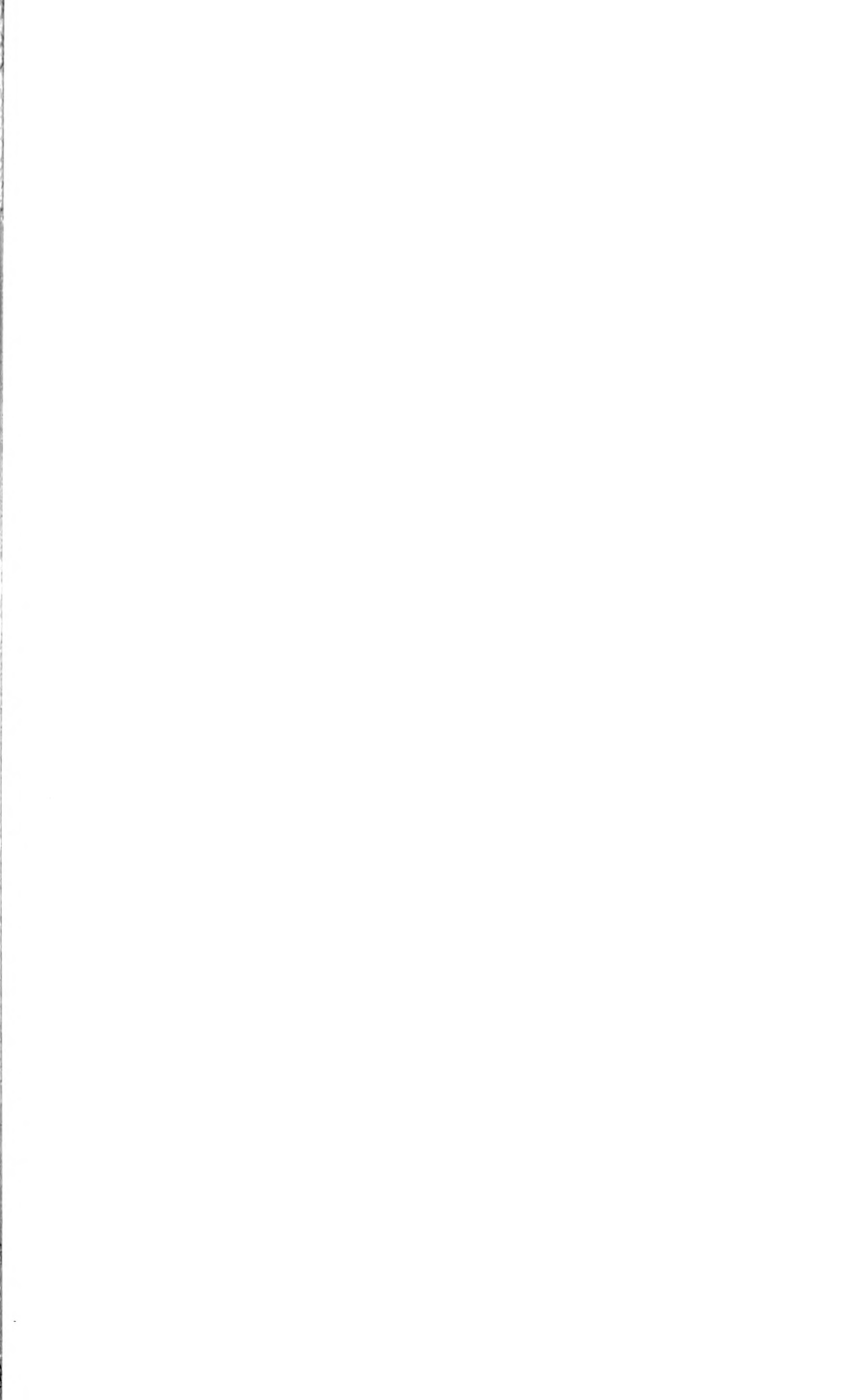
The ratio of release current to operate current depends on the armature residual, and if an adjustable residual screw is used some control of the release current can be achieved. If a small differential (i.e. difference between operate and release currents) is required, a large residual is necessary, and the armature travel may have to be increased.

The adjustment of a relay to definite operate and release currents thus becomes a matter of varying three interdependent parameters, viz. spring tension on to the armature, residual screw adjustment and armature travel. Inevitably a marginally adjusted relay is less sensitive than one in standard adjustment. Any attempt to combine sensitivity with marginal operation by using very light spring tensions and minute contact clearances usually results in unreliability.

Assembling relays

Even in a laboratory where relays are frequently needed it is not possible to stock complete units to suit all applications. It is then most convenient to stock the separate components of 3000 type relays and to assemble them as required.

A suitable selection of coils should first be provided. Resistances from 50 to 10,000 ohms will cover most requirements; a number of twin coil units are also desirable. Complete contact sets in useful combinations should be available in both left- and right-hand versions. Buffer blocks of the five different types are needed, together with armatures, yokes, mounting screws and coil nuts.



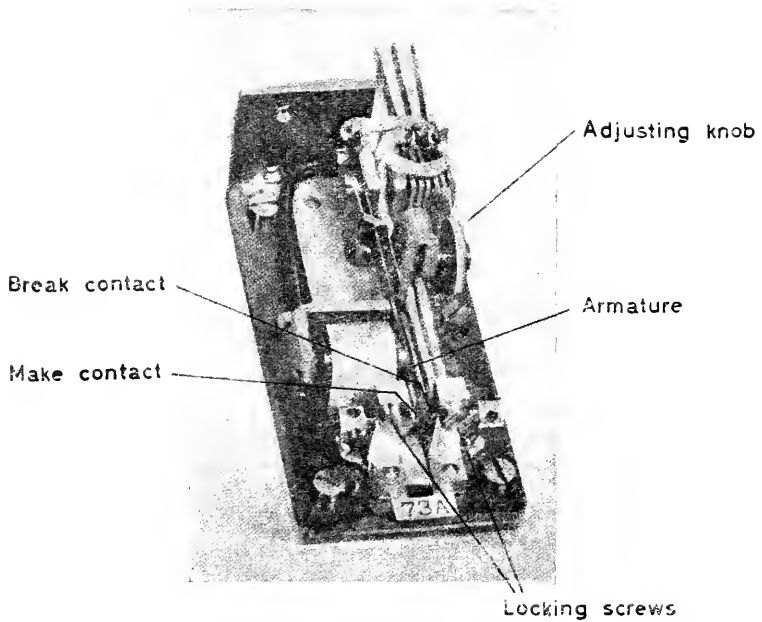


Figure 34.9 Siemens high-speed relay

HIGH-SPEED RELAY

The assembly of relays is straightforward, but certain points should be noted: (a) it is essential for all screws and nuts to be really tight; (b) the correct buffer block must be used; (c) a two-hole washer must be used under the buffer block retaining screws; (d) the correct length mounting screws must be used for contact sets and buffer blocks: short screws give inadequate strength, while long screws may penetrate the coil windings—the standard lengths of 6BA screws used in 3000 type relays are $\frac{1}{2}$, $\frac{3}{4}$, $\frac{3}{8}$ and $1\frac{1}{8}$ in.; (e) contact loads on the two sides of the buffer block should be approximately equal; (f) the relay must be adjusted after assembly.

It is strongly recommended that complete contact sets should never be assembled from individual springs and spacers. It is possible in this way inadvertently to construct contact sets which cannot be operated with a normal armature travel, which do not mate up to standard buffer blocks, which have inter-contact shorts through operating pins, and which cannot ever be adjusted for reliable operation.

HIGH-SPEED RELAY

When operate and release lags of less than 5 ms are required it is necessary to use a specially designed relay. Such a relay is shown in *Figure 34.9*; the operate and release lags have been reduced to about 1 ms.

Single or twin coil relays are available with resistances up to about 2,000 ohms. A single change-over contact is provided, using platinum as the contact material. Up to 0.5 amp may be carried by the contacts, but because of the small gap, operation on d.c. voltages over 50 is inadvisable unless particularly efficient spark-quench circuits are used.

A high-speed relay is adjusted as follows. Both contact screws are slackened back and the 'make' screw is adjusted so that it just touches the lever contact when the armature is held down on to the pole pieces. The make screw is advanced $\frac{1}{4}$ – $\frac{1}{2}$ turn to set the residual gap. The break contact screw is now adjusted to give a contact gap of 0.004 in. The spring is tensioned by turning the knurled knob until 20 g wt is needed at the armature tip to open the break contacts.

If a high-speed relay is to be operated in a circuit (such as a chopper amplifier) where contact bounce must at all costs be avoided, it is desirable to carry out the adjustment while the relay is being repetitively operated, and with a voltage interrupted by the contacts displayed on an oscillograph. To minimize contact bounce it is sometimes necessary to operate with higher armature tension than 20 g wt. High-speed relays can be operated as choppers satisfactorily up to about 100 c/s, although for a reasonable duty cycle a square wave driving current is advisable.

The power required by the Siemens high-speed relay for various conditions is shown below.

Operate	200 mW	Hold	80 mW
Non-operate	110 mW	Release	60 mW

For reliable operation, the usual factors of safety (page 517) should be applied.

Since twin coil relays have two separate coils on different pole-pieces, these

powers apply to operation by either coil alone. If, however, power is applied equally to the two coils, the total power requirement of the relay is only half the value tabulated above.

It is advisable to keep the mean dissipation in the coil below 1.5 W to prevent overheating.

POLARIZED RELAY

By incorporating a permanent magnet into the magnetic circuit a relay of high sensitivity and speed can be constructed. In addition, the direction of movement of the armature becomes sensitive to the direction of the operating current; such relays are therefore known as 'polarized relays'.

The design of a typical polarized relay is illustrated in *Figure 34.10*. The

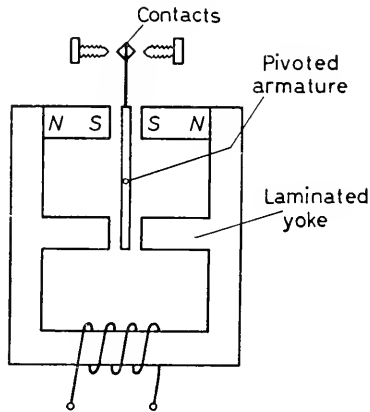


Figure 34.10 Polarized relay

exact arrangement of the armature control spring depends on the required resting position of the armature when the relay is un-energized. In a 'one-side stable' relay the armature normally rests against one fixed contact, changing over to the other only if an operating current of the correct polarity is supplied. The armature of an 'each-side stable' relay will rest stably on either fixed contact; when operated by a current of the appropriate polarity the armature snaps over to the other fixed contact. 'Centre stable' relays have no bias on the armature, which may rest anywhere between the fixed contacts. Some form of electrical bias is then necessary to centralize the armature, and in extreme cases to prevent it resting lightly on one or other fixed contact. In another type of centre-stable polarized relay the armature is positively constrained in the centre of the contact gap; this is achieved only at the expense of some loss of sensitivity.

The operating power for a polarized relay is between 0.2 and 2 mW, according to the design. Coils usually have a maximum dissipation of 1–2 W. Multiple coil windings can be provided. The contacts, which are normally a palladium-copper alloy, can carry up to 10 W with overall maxima of 1 amp and 200 V. Since small contact clearances (a few thousandths of an inch) are used it is essential that efficient spark-suppression circuits should be provided. The time of transit of the armature between the fixed contacts is about 1 ms.

OTHER DESIGNS OF RELAY

Polarized relays are particularly suitable for use with 'chopper' amplifier circuits (Chapter 39). For this application special types are available with platinum contacts for low noise, and with screening between the coil and the contacts to minimize spurious pick-up. Chopper relays should be of the each-side stable type, giving good contact pressure with high speed of operation; chopping frequencies of up to 300 c/s can be used. A chopper relay should be adjusted when running in its final circuit by observing the contact waveform on an oscillograph.

A particularly sensitive form of polarized relay is based on a moving-coil meter movement (Chapters 32 and 33) with contacts instead of a pointer; such a relay may operate on less than $1 \mu\text{W}$.

If an appreciable load is to be switched, an intermediate relay is essential as the contacts of moving-coil relays have a very low power rating. Moving-coil relays are both delicate and expensive, and in most applications can be replaced by a transistor and an ordinary relay (see Chapter 29). There are, however, a few purposes for which the use of a moving-coil relay is essential.

Ordinary non-polarized relays can be made to operate only on current of one polarity with the aid of a rectifier (*Figure 34.11*). Naturally none of the

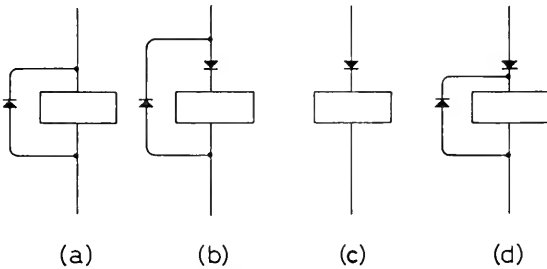


Figure 34.11 Rectifier-polarized relay: (a) and (b) constant-current feed; (c) and (d) constant-voltage feed

other advantages of the polarized relay is thereby obtained. Conversely, a polarized relay can be operated by a current of either polarity if it is connected in the circuit of *Figure 34.12*.

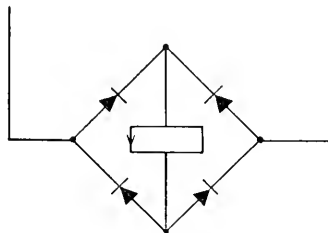
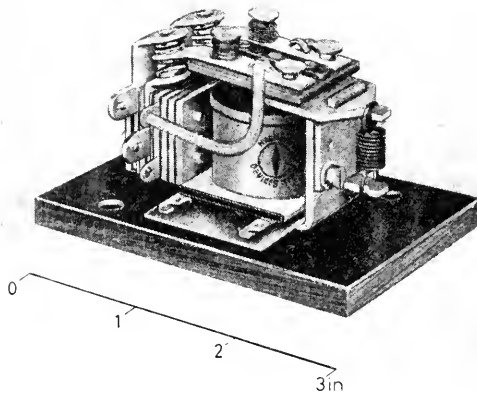


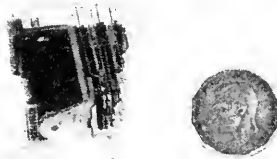
Figure 34.12 Operation of polarized relay on current of either polarity

OTHER DESIGNS OF RELAY

Many different designs of relay exist besides those already discussed. In all cases the fundamentals described for the P.O. 3000 type apply, although



(a)



(b)

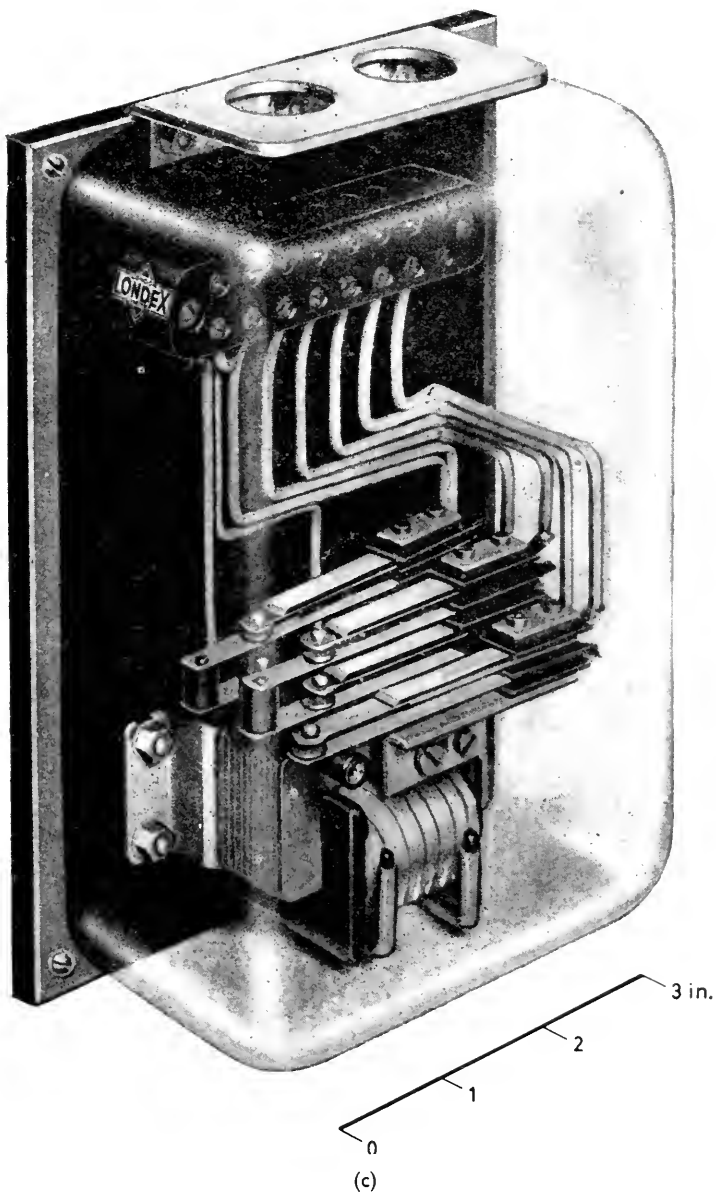


Figure 34.13 Examples of relays. (a) by courtesy of Magnetic Devices, Ltd.;
(c) by courtesy of Londex, Ltd.

there are many differences in mechanical construction. In this section these mechanical details will be discussed and illustrated by examples of commercially available relays.

Many relays use a clapper-type armature, either hinged or pivoted at one end (e.g. *Figure 34.13a*). The contact spring tension is then not usually called upon to restore the armature, this function being performed by a separate return spring. For very heavy-duty relays a solenoid type action is often used, with a plunger moving into a hollow coil.

Instead of a buffer block it is quite common practice to use stiff buffer springs behind the appropriate contact springs (*Figures 34.13b* and *c*). In another design (*Figure 34.13a*) the moving contacts are loosely attached to the armature, the contact pressure being provided by coil springs. Flexible 'pigtailed' connect the contacts to terminal tags.

Miniature relays (*Figure 34.13b*) are available with designs somewhat similar to a telephone relay, although the number of contact sets is usually more limited.

With all the types of relay illustrated it is essential that the manufacturer's data be consulted for details of operating power, contact rating, etc.

Relays for a.c. operation

Relays whose coils are energized with a.c. generally have laminated yokes and armatures. In addition, a copper ring is often embedded in the pole face to prevent the magnetic flux ever dropping to zero. Otherwise the construction of the relay follows conventional practice. An example of an a.c. relay is given in *Figure 34.13c*.

Any d.c. relay can be adapted to work from a.c. when combined with a metal rectifier. One of the circuits shown in *Figure 34.14* may be used.

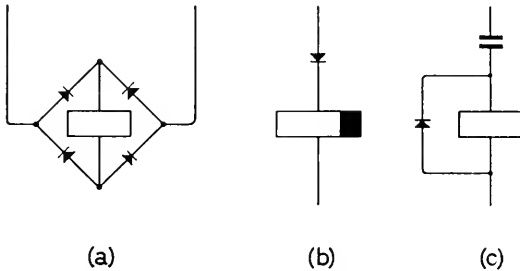


Figure 34.14 Operation of ordinary relay on a.c.

Mercury relays

For very high currents mercury switches (*Figure 34.15*) have many advantages over solid contacts. Thus with a tube about 1 in. long, 250 V 15 amps can safely be switched by a tilting movement of about 10 degrees. Various contact combinations—make, break, change-over and make-before-break—are available, and time-lags of up to three minutes can be arranged by using constricted tubes.

While complete assemblies of relay mechanism and mercury switch are commercially available, it is quite possible to mount a tube on a standard relay frame, such as the 3000 type. The mounting should be slightly resilient

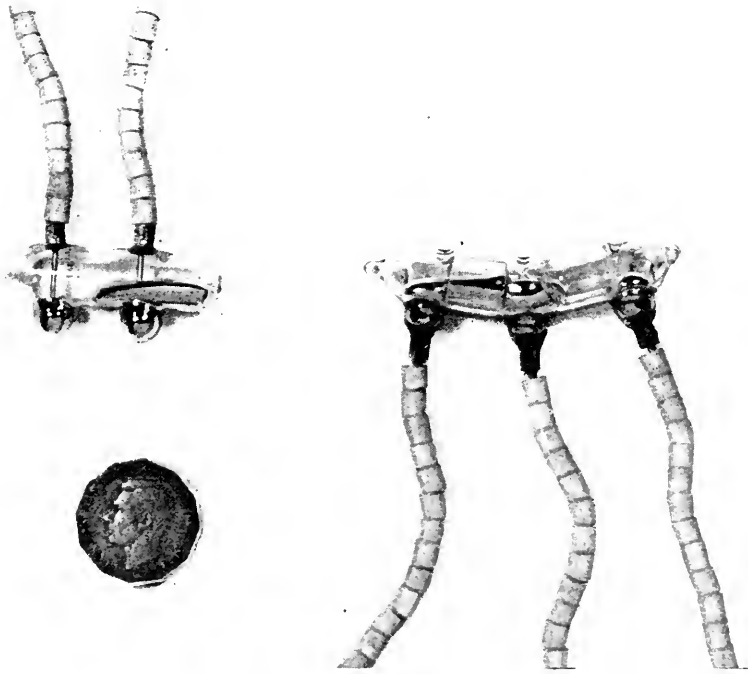


Figure 34.15 Mercury switches



OTHER DESIGNS OF RELAY

to avoid too violent a mechanical shock to the tube as the relay operates. A resilient mounting also permits slight expansion of the glass as the switch warms up under load.

If switch operation must be positive and free from 'bounce', some elastic connection between the relay armature and the switch is also desirable to prevent splashing of the mercury.

Mercury switches are also obtainable with the relay armature enclosed within the glass. An external electromagnet is then needed to move the armature and hence produce contact operation. A plunger switch of this type mounted within a solenoid can control currents up to 60 amps.

All mercury relays are orientation-sensitive, and it is essential that they be fixed at the correct angle on a rigid mounting.

For further details of the types of switch available, manufacturers' literature should be consulted.

Thermal relays

Electromagnetic action is not the only method of achieving mechanical movement in a relay: expansion caused by electrical heating is the basis of the operation of thermal relays.

In the thermal delay relay (Figure 34.16) the operating current is passed through a heating resistance wound on a bimetallic strip; when hot the

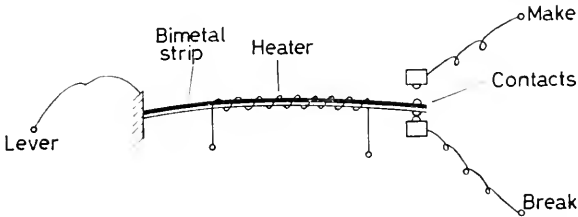


Figure 34.16 Thermal delay relay shown diagrammatically

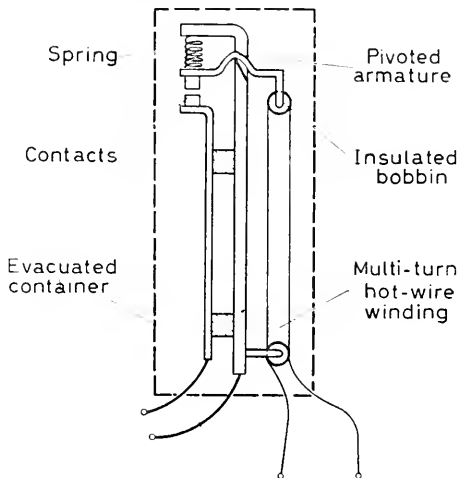


Figure 34.17 Hot-wire vacuum switch shown diagrammatically

strip bends and operates the contacts. A snap action is usually arranged, either by magnetizing the contacts or by using a bowed spring.

The operating time depends on ambient temperature and upon the immediate past history of operation of the relay. Operate and release lags of the order of one minute are possible with a relay of this type, but the accuracy of the timing is low (of the order of ± 30 per cent). The operating power required is several watts.

Thermal delay relays are most useful for the control of sequential circuit switching (e.g. switch on heaters—wait—switch on HT) where the disadvantages of low accuracy and high power consumption are of little significance.

Another device dependent on thermal expansion is the hot-wire vacuum switch (*Figure 34.17*). With a relay of this type currents of up to 25 amps can be switched, the operating power needed being only a few watts. Hot-wire switches can only be used when the delays of several seconds inherent in their operation can be tolerated.

RELAY CIRCUITS

In this section certain selected topics in the design of relay circuits are discussed. The uses of relays are so diverse that it is impossible to do more than illustrate by these examples some of the techniques available.

Circuits designed by telephone engineers tend to follow certain stereotyped conventions, based on the conservatism which is essential in the design of reliable telephone systems. However, the combination of relays with valves and transistors may produce circuits which although unorthodox both to telephone and electronic engineers are nevertheless highly efficient and elegant.

Spark-quench circuits

When an inductive circuit is interrupted the energy stored in the inductance, unless otherwise dissipated, will appear as a spark at the interrupting contacts. The contacts will eventually become pitted by the repeated sparking and unreliable operation will result; furthermore, radio interference will be produced, transmitting impulses to nearby high-gain amplifiers. Indeed, a relay operated from a sensitive amplifier has been known to go into a state of continuous oscillation due to feedback of spark interference into the amplifier input.

It is rarely necessary to use spark-suppression circuits on contacts controlling other relays, but if uniselectors, motors or other heavy and highly inductive loads are switched, some form of spark quenching is essential. *Figure 34.18* shows several possible circuits which provide a path for the stored inductive energy. The most generally useful circuit is that of *Figure 34.18c* in which a non-linear resistor (e.g. Metrosil, Atmite) is used. Under normal conditions little current flows through the non-linear resistance, but the high voltages developed at the interruption of the inductive circuit cause its resistance to fall, and large momentary currents pass. The voltage across the contacts is therefore kept below the value at which sparking will occur.

All spark-suppression circuits, by providing low resistance paths for circulating currents, delay the collapse of magnetic flux in the inductive load;

RELAY CIRCUITS

therefore a relay will have a significantly greater release lag if its controlling contacts are fitted with a spark-quench circuit.

Conventional spark-quench circuits sometimes fail to suppress arcing at contacts breaking large direct currents. In these rare cases a small horse-shoe

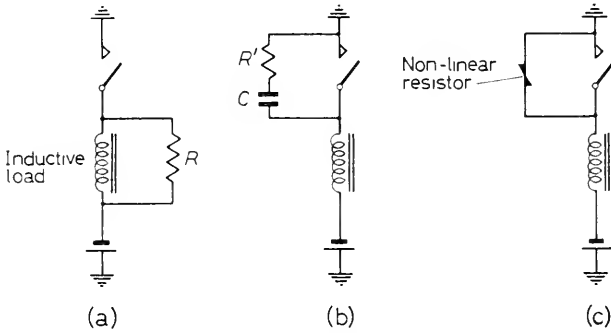


Figure 34.18 Spark-quench circuits

Typical values			
Load	R	R'	C
75 Ω uniselector magnet	350 Ω	200 Ω	1 μF
1,000 Ω relay	10,000 Ω	5,000 Ω	0.05 μF

magnet can be mounted so as to give a strong magnetic field at right angles to the arc path. The resulting electromagnetic forces 'blow out' the arc immediately after it forms.

Self-locking circuits

If the relay in Figure 34.19a is operated by closing the key *L*, it will thereafter remain operated via *A1* even when *L* is released. The relay is said to be

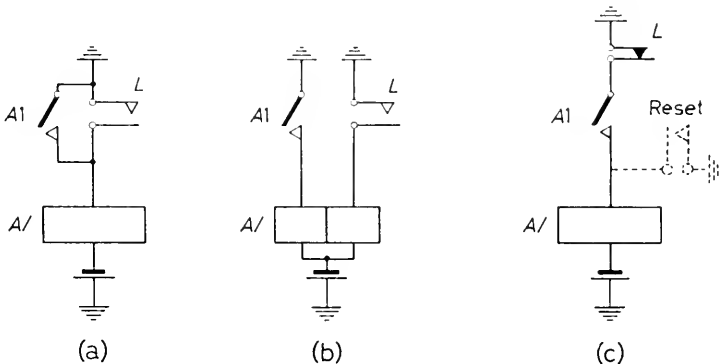


Figure 34.19 Self-locking circuits

'held' via the 'hold contact' *A1*. The relay can of course be released if a series break contact is provided. Another common arrangement is shown in Figure 34.19b; a separate 'hold winding' is provided. This type of circuit is

needed when, for example, two relays are operated in parallel by L , but only one is required to hold.

The inverse circuit is shown in *Figure 34.19c*. Here the relay is released if L is opened and remains released even if L re-closes. The circuit may be re-set by the make contact shown dotted.

External control of operate and release lags

In a previous section it has been shown that the operate and release lags of a relay depend to a great extent on the external circuit configuration. It

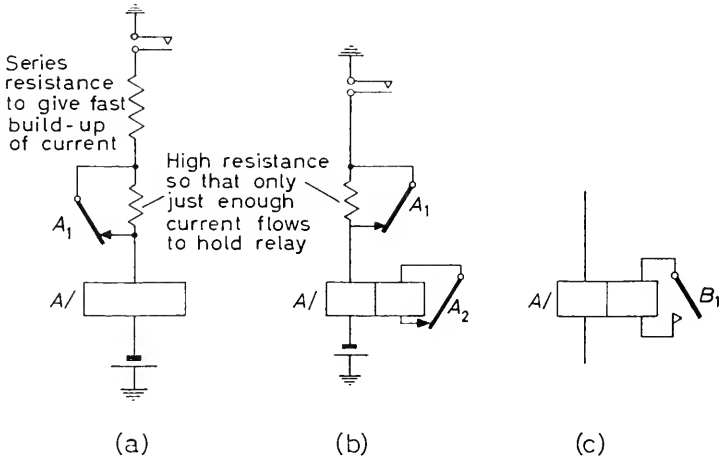


Figure 34.20 External control of relay timing: (a) fast operate, fast release; (b) slow operate, fast release; (c) lag of relay A depends on operation of relay B

will be seen from *Table 1* that some of the conditions are mutually incompatible; thus a relay for fast operation must be provided with an excess of operating power, while for fast release just sufficient power should be used. Fortunately the operation of the relay can be used to change the external conditions with suitable contacts, so that it is possible to meet both operating and release requirements.

If the second coil on a twin coil relay is short-circuited, an effect similar to that of a slug is produced, and the relay action is slowed down. The coil can be short-circuited by one of the relay's own contacts.

Some examples of the use of relay contacts to give specified operate and release conditions are shown in *Figure 34.20*: the contacts need not, of course, belong to the relay whose timing is being controlled. Thus the speed of operation of a relay can depend on whether a second relay is operated or released (*Figure 34.20c*).

It is not possible by any circuits to this type to exceed the delay available from a slugged relay (operate, 100 ms; release, 500 ms). For greater delays a resistance-capacitance circuit can be used, as shown in *Figure 34.21a*. Still longer delays can be obtained with a twin coil relay, as in *Figure 34.21b*. Again the capacitor can be switched by one of the relay's own contacts so that operate and release lags can be controlled independently.

RELAY CIRCUITS

Even with 'capacitance slugging', delays of much more than ten seconds are not possible. By using thermally sensitive resistors delays of up to a minute or so may be obtained. The duration of the delay will not, however, be accurate, but will depend on ambient temperature. In addition, an adequate time must be left between operations for the thermal element to revert

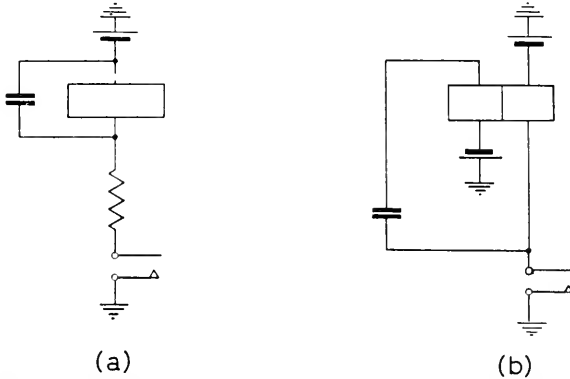


Figure 34.21 Capacitance delay circuits: (a) single coil relay; (b) twin coil relay

to room temperature. Figure 34.22 shows a group of circuits using the large and cheap thermistor elements of the type used as current limiters (e.g. Brimistors).

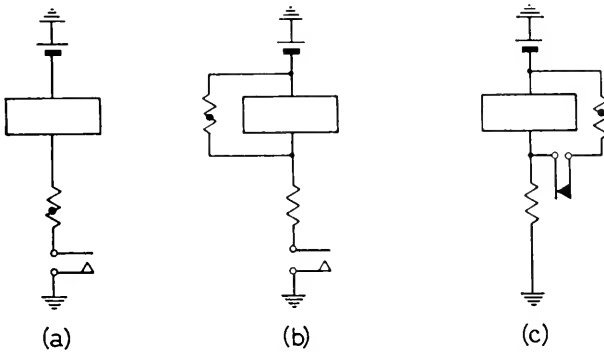


Figure 34.22 Thermistor delay circuits: (a) slow operate, fast release; (b) fast operate, then after delay, release; (c) fast operate, slow release

Counting circuits

Relay circuits can be used to 'count down' impulses so that an output is provided only at every second (or third, etc.) input impulse. In this section some of these circuits are described, both for their inherent usefulness and also as examples of relay circuit design. For other examples of scales-of-two, a comprehensive review of scaling circuits by Barnes² should be consulted.

Figure 34.23a shows a typical scale-of-two circuit. When the input key *L* is closed, relay *A* operates through the make-before-break contact *A1*; it

RELAYS AND RELATED MECHANISMS

then holds through the operation of the same contact. $A1$ also provides an operate path for B , which however has its other coil connection earthed via $B1$ and L . When L is released, B operates. The second operation of L short-circuits A via $B1$ and L , and it releases. B is now held via $A1$ and L , and upon the opening of L it releases. The sequence of relay operations is shown in *Figure 34.23b*. An auxiliary contact on either A or B thus closes once per two operations of L .

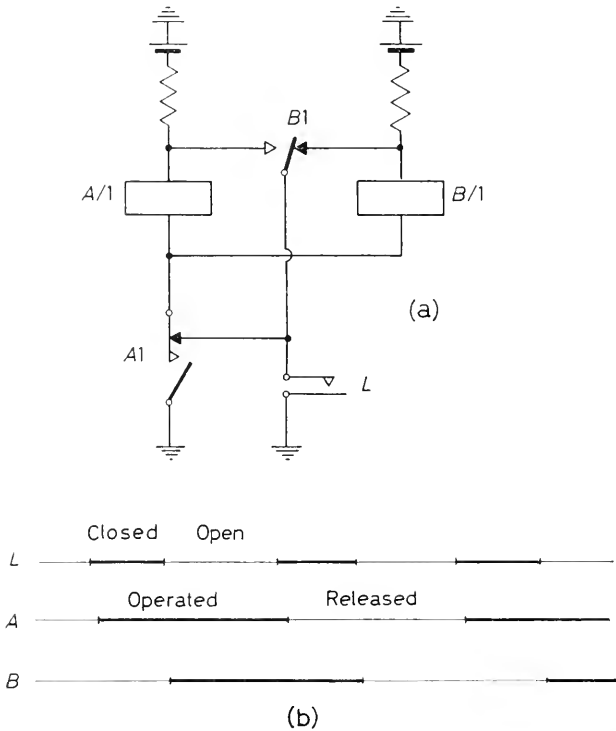


Figure 34.23 Scale-of-two: (a) circuit; (b) contact sequence diagram

A circuit with similar performance is shown in *Figure 34.24*. The make-before-break contact is dispensed with, but the operating contact L is no longer earthy. The resistor R serves to protect the supply if contact $B1$ bridges it due to faulty adjustment or other failure.

Both the circuits described rely on the releasing of a relay by short-circuiting. This inevitably makes their operation slow, due to the large release lags inherent in this method. It is not possible to overcome this defect with a two-relay circuit using a simple make contact for input.

A circuit in which the release of both relays is achieved by open-circuiting is given in *Figure 34.25*. The sequence of operations is again shown by *Figure 34.23b*. Now the input is applied by a make-before-break key, which must however never dwell in the half-operated condition. If this cannot be prevented by its mechanical design, an auxiliary relay C may have to be provided.

RELAY CIRCUITS

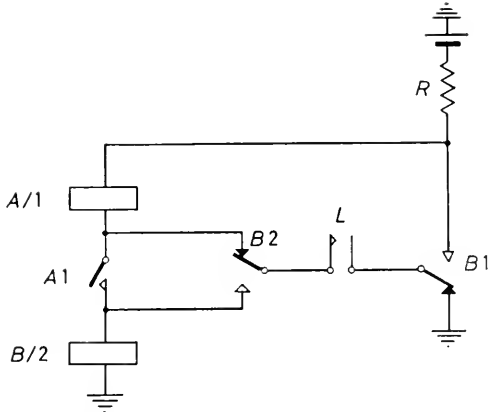


Figure 34.24 Alternative scale-of-two circuit

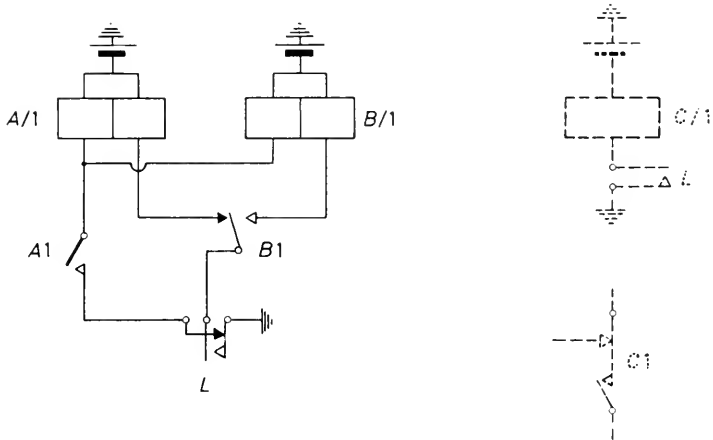


Figure 34.25 Fast scale-of-two circuit

With three relays it is possible to design very simple and fast scalers. Figure 34.26 shows a design in which only one contact per relay is needed; it is therefore suitable for use with high-speed relays. For the highest speed of operation each relay should be fed from a high-impedance source; for a 145 ohm relay, a 50 V supply with a 600 ohm series resistor is suitable. Pulse rates of up to 100 per second can then be handled.

When it is required to count down by more than two, a series of n scales-of-two can be cascaded to give a division by 2^n . Alternatively, ring circuits of the type shown in Figure 34.27 can be used. Although illustrated as a scale of three the circuit can be extended to any greater number of relays. Contacts B4, C4 ensure that when the supply is switched on relay A operates and holds; if high speed of operation is not essential these contacts may be replaced by a large capacitor.

RELAYS AND RELATED MECHANISMS

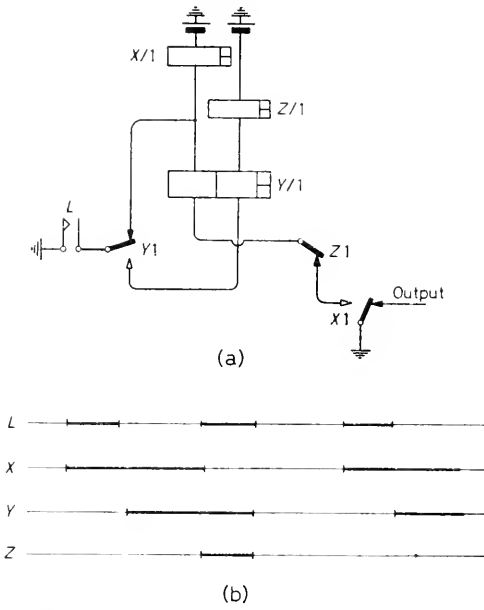


Figure 34.26 High-speed relay scale-of-two: (a) circuit; (b) contact sequence diagram

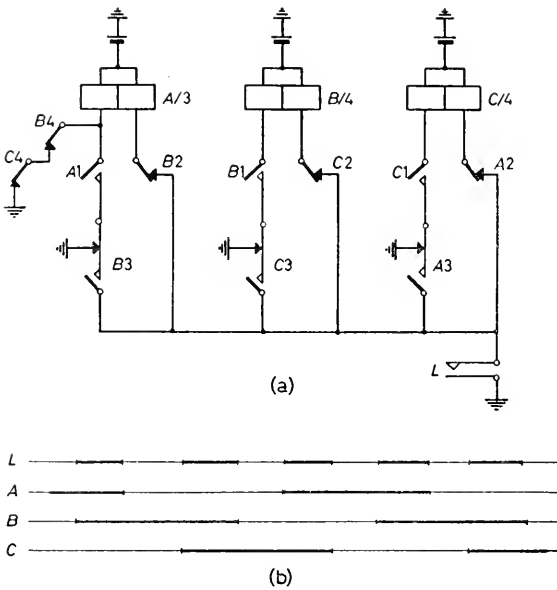


Figure 34.27 Ring-of-three: (a) circuit; (b) contact sequence diagram

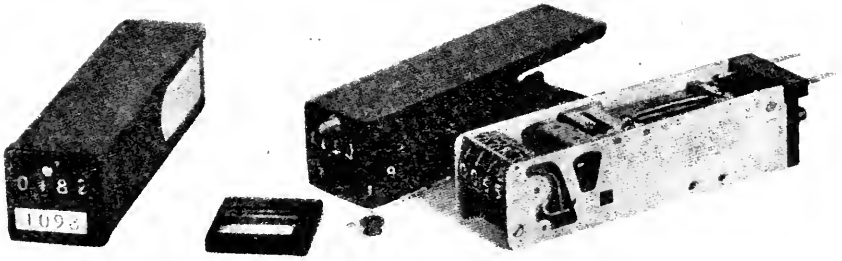


Figure 34.28 P.O. Electromagnetic counter

(By courtesy of J. Atkinson and Sir Isaac Pitman and Sons, Ltd., from *Telephony*, 1948)

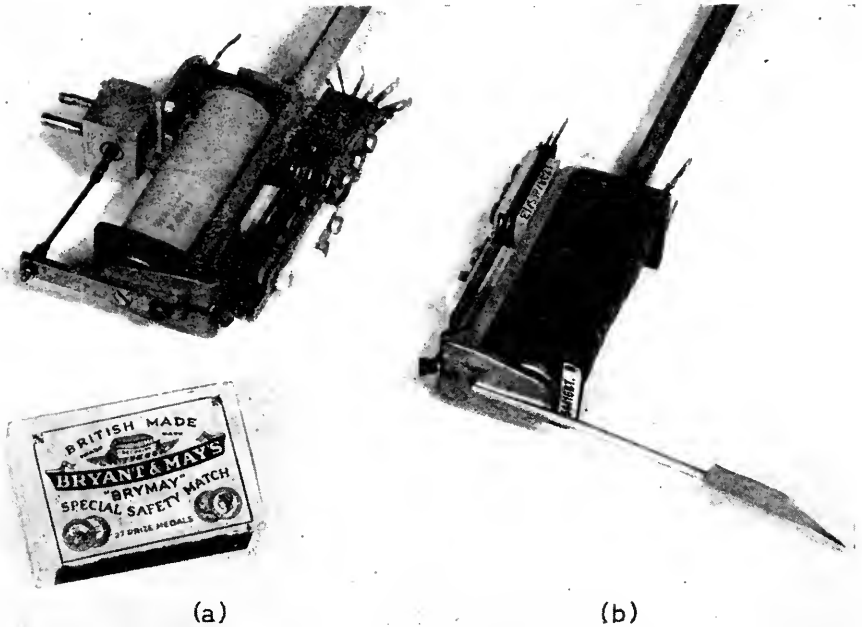


Figure 34.29 Two mechanisms using 3000 type relay frame: (a) pneumatic valve; (b) smoked-drum pen

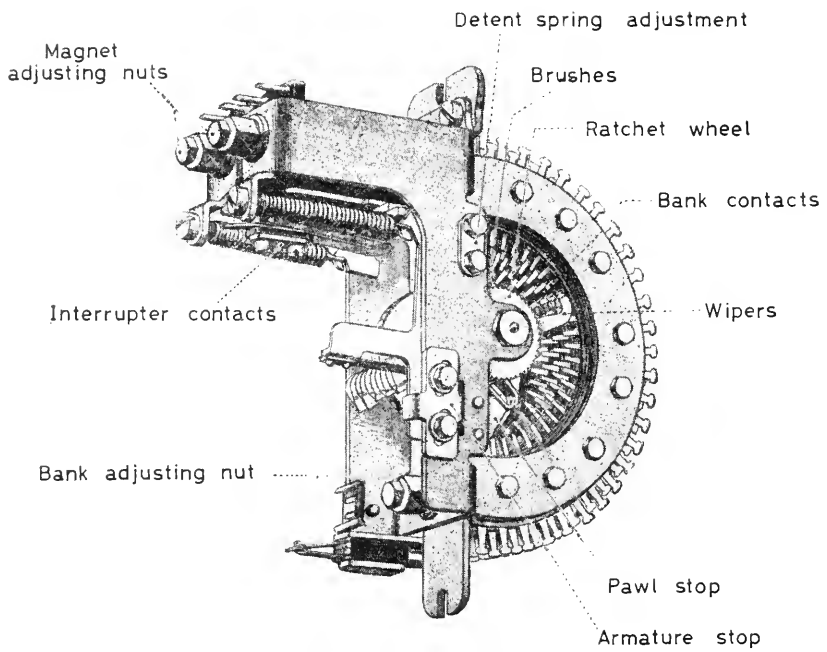


Figure 34.30 Uniselector

(By courtesy of J. Atkinson and Sir Isaac Pitman and Sons, Ltd. from *Telephony*, 1948)

UNISELECTORS

RELAY-TYPE MECHANISMS

The mechanical construction of certain mechanisms is based on that of relays, and since they are commonly used in combination with relays they are briefly described in this section.

Electromagnetic counter

The standard Post Office electromagnetic counter or message register is shown in *Figure 34.28*. It counts up to 9999 but has no method of re-setting. An operating power of about 300 mW is required, and impulses can be registered at speeds up to five per sec.

Other designs of counter available include types with manual or electrical reset, and types which print their readings on a paper strip each time a second coil is energized.

Adaptations of the 3000 type relay

The mechanical movement of most relays can be adapted to purposes other than the operation of contacts. Two mechanisms based on the 3000 type relay are shown in *Figure 34.29*. The operating power required for devices of this type must be determined experimentally. It should be remembered that because of magnetic saturation it is not possible to achieve armature tractive forces of more than about 800 g wt; this will require an operating power of about 2 W.

Modifications to relays do not usually prevent their use in their original form: thus both the devices of *Figure 34.29* carry auxiliary contacts controlling other circuits.

UNISELECTORS

The uniselector is an electrically driven rotary selector switch which has many applications for counting impulses and carrying out complicated sequential switching operations.

The construction of a uniselector is shown in *Figure 34.30*. Electrically it provides a 25-way switch of up to eight poles, the switches being either of the break-before-make ('non-bridging wiper') type, or make-before-break ('bridging wiper'). Non-bridging wipers have narrow contact tips which can only touch one bank contact at a time, while bridging wipers have broader tips which short-circuit two adjacent contacts while moving. By cutting off opposite ends of alternate wipers 50-way switches can be produced; thus contacts 1–25 will be on level 1 while contacts 26–50 will be on level 2. Wipers 1 and 2 are, of course, commoned.

Uniselectors require about 30 W to operate; they can be driven at up to about 25 steps per second. Interrupter contacts are provided so that the mechanism can be arranged to 'self-step' and rotate continuously. Spark suppression is essential for reliable operation.

The adjustment of a uniselector is somewhat complicated but essential for reliable operation. The steps are as follows:

- (1) Check that the pressure of each wiper tip on the bank contacts is about 30 g wt. Adjust if necessary with the two-pin relay adjusting tool.

RELAYS AND RELATED MECHANISMS

(2) Check that the pressure of the brush wires on the slip-rings is about 35 g wt. Adjust if necessary.

(3) Set the wipers to contact 1, and adjust the armature stop until the tips of non-bridging wipers lie just behind the leading edge of the bank contacts.

(4) Set the pawl stop just touching the pawl.

(5) Set the detent spring so that it fits snugly into the teeth of the ratchet wheel.

(6) Set the wipers to contact 25, and turn the bank adjusting nut until the wipers again lie correctly relative to the bank contacts. Lock the nut.

(7) Adjust both magnet adjusting nuts so that the operation of the armature moves the pawl just more than one tooth, and so that both pole-pieces are at the same height. Lock the adjusting nuts.

(8) Ensure that the armature springs have sufficient tension to return the armature positively.

(9) Connect a self-stepping circuit, and adjust the interrupter contacts for steady rotation.

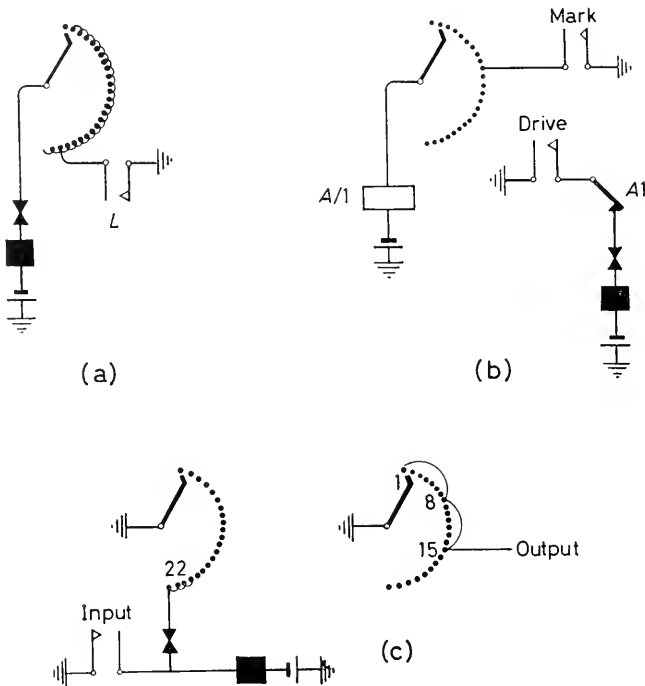


Figure 34.31 Uniselectors circuits: (a) homing; (b) drive to marked position; (c) count down by seven

Uniselectors circuits

In this section three simple circuits are shown to illustrate the bases of the common methods of unisector operation.

REFERENCES

In *Figure 34.31a*, closing of key *L* makes the uniselector 'home' to a defined position (usually contact 1). *Figure 34.31b* shows a circuit whereby the uniselector drives to any 'marked' position and then stops. When it is desired to 'count-down' by a number other than 25 a circuit similar to *Figure 34.31c* may be used.

REFERENCES

- ¹ ATKINSON, J. *Telephony*, Chap. VI. Pitman. 1948
- ² BARNES, R. C. M. *Electron. Engng.* 26 (1954) 432, 493

GLASS MICROCAPILLARY ELECTRODES USED FOR MEASURING POTENTIAL IN LIVING TISSUES

D. W. KENNARD

MICROCAPILLARY ELECTRODES FOR INTRACELLULAR RECORDING BY TRANSVERSE MEMBRANE PUNCTURE

Electrical potentials and currents may be detected about most biological tissues. Those of greatest physiological interest are due to work performed by living tissues in separating the electrically charged ionic particles. The electrical energy required to do this work can be stored in tissues such as nerve and muscle which form concentration cells, from which it can be liberated without the aid of metabolic activity. Tissues may be regarded as batteries which pass current through an external circuit composed of the saline medium of the living organism.

To study the electrical activity of individual cells an electrode can be introduced into the interior and the potential difference (P.D.) across the cell walls determined. If the cell is large enough the electrode can be introduced through a cut end, as was done by Hodgkin and Huxley¹ and Curtis and Cole², using the giant nerve fibre of the squid. They found that by introducing a capillary electrode of $60\ \mu$ diameter for a distance of 2–3 cm undamaged portions of the cell could be studied.

For most cells the cell wall has to be punctured by the electrode itself, which is made as small as possible to minimize damage. The first successful punctures of this type were made by Ling and Gerard³ using electrodes of about $0.5\ \mu$ diameter in the striated muscle fibres of the frog. Electrodes for intracellular work are required to record both steady potential levels and rapid fluctuations of voltage. The simplest method of establishing connection between a tissue and a measuring device is to use a suitable metallic conductor. Such a connection, though useful for recording fluctuations of voltage, introduces difficulties in measuring steady levels. When any metal is immersed in a saline medium a P.D. is established between it and the fluid; the actual values depend on the nature of the metal and the fluid. For example the tungsten electrode potential depends on the concentration of hydrogen ions present (Britton⁴) and has been used to measure intracellular pH (Caldwell⁵). The potential of a silver electrode depends on the halide ions present and has been used to estimate chloride ion.

During the course of an experiment, however, the composition of tissue fluids may vary, causing potential changes. To ensure stable electrode voltages the electrode environment must be constant, and this can be achieved

by placing the metal electrode away from the tissue, where the chemical environment is not subject to variation.

The glass microcapillary electrode serves as an electrolyte bridge between the tissues and the electrode surface; a fluid junction is established at the tip within the tissues. When two different electrolyte solutions form a junction a P.D. is established across it, due to the different mobilities and concentration of ions present, so that the choice of the medium is important. Thus at a sodium chloride junction the smaller Cl^- ion will tend to diffuse faster than the Na^+ ion, causing a P.D. to be established. A P.D. of over 3 mV is found between two solutions of sodium chloride with concentration ratio 1 : 10. However with potassium chloride solutions, in which the K^+ and Cl^- ions have similar mobilities, the same concentration ratio gives a P.D. of only 0.4 mV. Concentrated solutions of potassium chloride—usually 3 M—are used to establish fluid junctions in order to reduce the diffusion potential due to the other ions present. As a result of their high concentration the junction currents are largely carried by K^+ and Cl^- ions, thus reducing the potential. When 3 M potassium chloride solutions are employed the ionic content of small cells may be altered significantly by the diffusion from the tip. It was calculated by Nastuk and Hodgkin that 6×10^{-14} mole per sec would diffuse from the tip of a sample electrode, of 0.4μ outside diameter. This quantity would cause negligible change in the constitution of a muscle fibre, but could appreciably alter the content of a neurone (Eccles⁶).

Capillary electrodes of such small dimensions possess a high resistance. Nastuk and Hodgkin⁷ found that the resistance of their electrodes filled with isotonic potassium chloride was 5 to 7 times greater than when filled with 3 M solution. It is advantageous therefore to use such concentrated solutions to reduce the resistance to a minimum. Nevertheless, the intracellular electrodes generally used have resistances not less than 5 M Ω when measured in Ringer solution, and range up to about 100 M Ω . When measuring potentials with such electrodes the input resistance of the amplifier must be many times greater than that of the electrode. If R_1 represents the resistance of the electrode and tissue, R_2 the input resistance of the amplifier, and E the voltage applied, the potential recorded will be $E \cdot R_2 / (R_1 + R_2)$. Thus even when the ratio $R_1 : R_2 = 1 : 10$ the recorded potential will be $E \times 0.91$. It is therefore necessary that the input resistance of the amplifier should be not less than about 10^{10} M Ω , in order to measure the applied voltage.

Construction of glass microcapillary electrodes

Electrodes for intracellular studies by membrane puncture are usually required to have tips of almost submicroscopic dimension, often less than 0.5μ external diameter. As a result their true proportions can be decided only with the electron microscope. Nevertheless the tips have to be sufficiently strong, in some cases, to be introduced through layers of tissue without breaking. The most commonly used glass for such electrodes is a hard borosilicate, usually either Pyrex or Phoenix, which has a high content of silica with B_2O_3 added to reduce the softening temperature, and the proportion of alkali oxides is small. In addition to its strength borosilicate glass

GLASS MICROCAPILLARY ELECTRODES

possesses other advantages both chemical and physical. It does not readily dissolve in water or give up its alkali. The volume resistivity is high, given by Morey⁸ as $3.1 \times 10^{14} \Omega \text{ cm}$, and its surface resistivity, due to adsorbed moisture, is better than many other glasses. Its dielectric strength can be several times that for soda glass, the breakdown voltage being approximately 1,000 kV per cm.

Hand-pulled electrodes—Microelectrodes of wide variety in shape and size can be produced by manual means.

Microelectrodes are usually made from tubing of 1–2 mm outside diameter, which is prepared from any larger size conveniently available. The lumen of the initial tube should be approximately two-thirds of the total diameter. A Pyrex tube 8–9 mm external diameter and 5–6 mm internal diameter is heated in an oxygen-gas flame and drawn down to about 1.5 mm, in lengths of 1–2 ft. The initial part of the draw is performed relatively quickly until the required diameter is reached, and a steady pull is then maintained to obtain a length of uniform thickness. Two in. lengths of this tubing are prepared.

For the final stages of drawing a very small oxygen-gas flame is used, which can be obtained from a gas-jet with a bore of approximately 0.3 mm diameter. A short length of tubing is held in both hands and 2 mm in the middle is heated to red heat indicating a temperature at which the glass has softened but not begun to collapse. The tube is then removed from the flame and rapidly pulled apart, when two electrodes should be formed.

Whilst drawing, both ends should be kept parallel to the direction of movement, as this will lead to the production of straight electrodes, and will prevent the occurrence of flattening. The length of the electrodes will be determined by the amount of glass heated, the temperature and the speed of the draw. The larger the flame the greater will be the amount of glass heated, thus increasing the length. The temperature of the glass, as indicated by the colour, should determine the time at which the draw is made. If too high, as when approaching white heat, a deliberate pause can be made to allow cooling to occur, and either the glass held stationary or slowly extended a few mm. This can in fact be done routinely as it allows more time for the exact drawing point to be estimated. Overheating however may lead to partial collapse of the tube, resulting in a reduction of the ratio of lumen to wall thickness. A variety of electrodes with varying shank length and diameter can be produced by varying the speed during the draw. Thus a slow pull with soft glass will produce a thick shank, while a rapid movement with cooler glass will yield a more sharply tapered and thinner shank.

A two-stage method can also be used. The tube of 1–2 mm diameter is first drawn down to form a short length about 0.3 mm thick; then, using a very small flame this tube is heated and drawn apart very rapidly. This method gives electrodes with a wide stem, thus eventually reducing the resistance to some extent; it requires, however, more skill.

If the tube at all stages is prevented from overheating and collapsing, the ratio of lumen to wall thickness remains constant throughout. It was found by direct measurement, from the end-on view, that this ratio varied less than 5 per cent commencing with the initial parent tube down to about

4 μ , and further it appears from electromicrographs of Nastuk and Hodgkin⁷ that this was also true at the very tip (Hodgkin⁹).

Mechanically produced microelectrodes—Several methods varying in degree of mechanization have been employed to make the production of electrodes less subject to variation in timing, temperature, strength and direction of pull. All these are difficult to maintain constant when drawing manually.

The simplest mechanical aid employs gravitational force for extension, and the electrodes are drawn in one or more stages. In the single-stage method the glass tube, 1–2 mm external diameter, is clamped in a vertical position between two rubber cushions and the coil, when used, lifted on a vertical slide to surround the tube. A weight is attached either by means of a clamp with rubber jaws (Templeton¹⁰) or more reliably to a hook formed from the end of the tube. Weights applied range up to 500 g, depending on the thickness of the glass. Heat is applied either with a gas flame or an electrically heated coil or wire. The gas jet can be made from a No. 20 Record needle or better from a glass tube with narrow tip, brought up by hand. The electrically heated coil is preferable as it can more readily be controlled. Coils of two or three turns can be used, between 3–5 mm in diameter and 3 mm long. Weale¹¹ used oxynickelnichrome wire of 2 Ω /in., with power from a 6 V battery in series with a rheostat; however wire between 0.3–0.5 mm diameter and a heavy duty variable transformer can also be used. When the tube has extended some 5 to 7 mm heat is turned off.

With this method capillary tubes between 3–30 μ can readily be made but greater difficulty is encountered in producing electrodes of less than 1 μ . The one advantage of this method as against hand drawing is that the electrode shank can be made collinear with the parent stem.

Electrodes drawn in several stages—Gravitational pull can be used to draw microelectrodes in several stages under controlled conditions. A method employing the de Fonbrume Microforge has been described by Brock, Coombs and Eccles¹² in which considerable care is taken in the manufacture of each electrode. This method is capable of yielding a much higher proportion of successful attempts than the single-stage draw. Pyrex tubing of 4–5 mm outside diameter is pulled by hand down to 0.15 mm with a length of at least 1 cm. A bead is formed at the tip and the tube mounted vertically with the tip down. The succeeding stages are observed with a microscope. The tip is coated with wax (Chatterton's compound or Picene) and a 300 mg weight attached by means of a thread and small hook. Using the electrically heated loop of wire the tube is drawn down to 50 μ diameter, when the weight is reduced to 30 mg and the final draw accomplished. The following modification can be used: a tube is first drawn by hand down to about 0.2 mm and as before a bead formed and thread with hook attached. A weight of 0.5–1 g is suspended and by means of a heated wire loop the tube is drawn under a dissecting microscope. The heating coil is mounted on a glass slide coated on one side with oil or Vaseline and held against a block of Perspex between thumb and fingers. Movements in the vertical plane are performed by the thumb holding the glass slide while those in the horizontal plane are carried out by sliding

GLASS MICROCAPILLARY ELECTRODES

the block on a flat surface. Essentially this arrangement merely serves to steady the hand. Although with the dissecting microscope the actual tip cannot be observed, this method is capable of giving good results with practice. With the multi-stage method the shape of each part of the electrode can be controlled, a considerable advantage for some experiments.

Other instruments for the production of microelectrodes utilizing gravitational force have been constructed in which the capillary is drawn horizontally in one stage (Kao¹³). On the whole it seems that if an elaborate construction is undertaken a flexible method of providing power should be employed.

Machines operated by springs—Many types of elastically operated micro-electrode pulling machines have been constructed, though few have been featured in publications. Three types are described here.

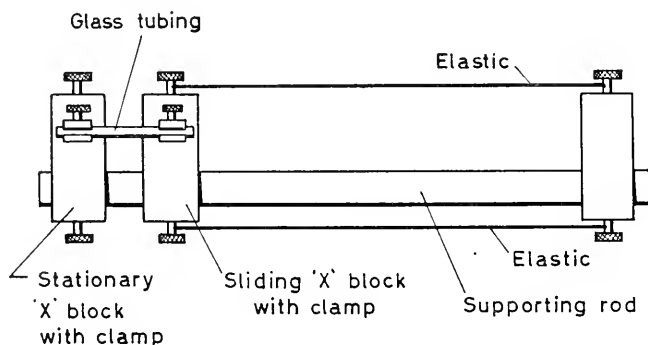


Figure 35.1 A simple microcapillary pulling device from a description by Crain¹⁴

The simplest mechanism is that in which one of a pair of clamps is mounted on a horizontal slide operated by a spring or elastic band (Crain¹⁴). Once the glass tube is clamped, the elastic is stretched to the required degree and heat applied either with a gas jet or electrically heated coil. An effective sliding arrangement can be made from an *X* block and a metal rod, with two elastic bands, one on either side of the rod (Figure 35.1). Various additional features may be provided such as a knock-down switch to cut off heating current, and a mechanical switch to bring in additional elastic bands.

A more intricate type of spring operated machine is shown diagrammatically in Figure 35.2. The action of this machine is similar to manual drawing. The glass tube is pulled apart and is simultaneously lifted away from the heating loop. Two toothed discs 7 in. diameter are engaged side by side, with rotatable clamps for glass tubing mounted near the periphery. A horizontal rod sliding through both clamps keeps them parallel at all positions, and heat is applied by an open *U* shaped loop of nichrome wire. The machine is operated by a vertically mounted coiled spring attached to a string which is wound round a grooved wheel mounted coaxially on one of the large toothed discs. The spring acts to unwind the string by rotating the discs, thus pulling the clamps apart. The pulling force can be varied by altering the length of the spring.

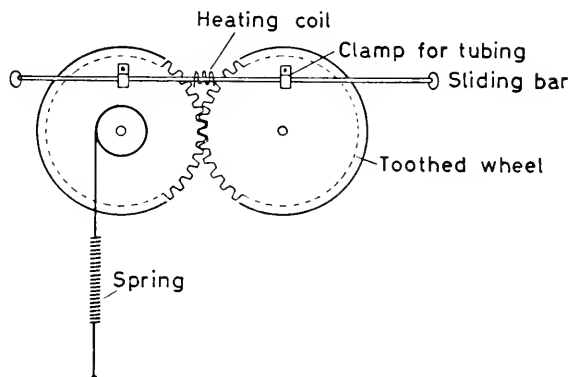


Figure 35.2 Schematic drawing of a microelectrode pulling machine

Another versatile mechanism, ascribed to J. W. Woodbury, is operated on a scissor principle. The tubing clamps are mounted at the ends of scissor action blades, as shown in Figure 35.3, operated by a spring causing them

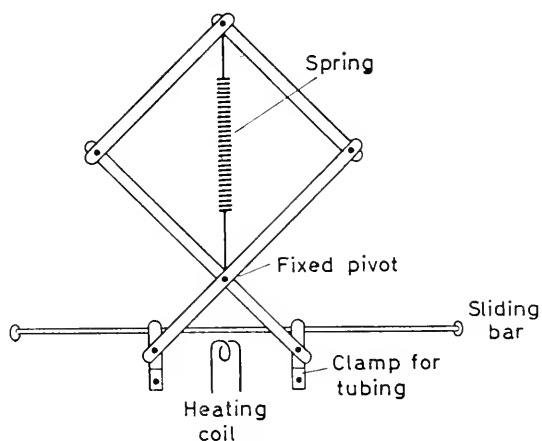


Figure 35.3 Schematic drawing of a microelectrode pulling machine

to open. A sliding bar passing through both clamps maintains them parallel throughout the action. The whole mechanism is pivoted and held at one point. A horizontal U shaped heating coil is used, or a gas jet. This machine can be readily constructed from simple parts, and also has the advantage that the act of pulling withdraws the glass from the heating element.

Machines operated by solenoids—A reliable and commonly used type of electrode pulling machine is based on the original pattern of Alexander and Nastuk¹⁵. This can give constant and repeatable results with a high proportion of successes, and facilities are available for varying the type of electrode produced. With this apparatus a two-stage draw is accomplished by a solenoid acting on a concentric iron plunger carrying the glass tubing. Power is derived from the mains supply.

GLASS MICROCAPILLARY ELECTRODES

A glass tube is passed through a platinum loop 3 mm in diameter and length and is held in a stationary clamp at one end and at the other in a clamp mounted on the plunger. The plunger is supported on ball races and is prevented from rotating. In its traverse it actuates two switches, one causing an increase in the pull of the solenoid and the other turning off the current to the platinum heating loop.

Once the tube is clamped in position and power supplied the action is automatic.

I Heater current is turned on. Solenoid exerts small tractive force.

II Glass softens and extends.

III (a) At pre-set point of extension heater current turned off. (b) At independently variable point force of solenoid greatly increased.

IV Solenoid current turned off.

The construction of this machine has been described in detail¹⁵ and will only be supplemented here with some notes, based largely on experience with a machine constructed in a modified form by R. H. Cook (*Plate 35.1*)

(1) Heater element. The clamps supporting the loop on to the heat resistant insulating block should be of minimal size so as to reduce their heat capacity, which otherwise prolongs the heating time. The element being in a vulnerable position, and liable to damage, can have a guard placed close on one side attached to the stationary clamp. This acts as a guide when fitting the tubing to be drawn (*Plate 35.2*). When required the loop is reformed around a 3 mm rod with pointed ends supported by the clamps.

(2) It should be possible to control and re-set the various parameters to any determined position, so that systematic trials can be made and different types of electrodes drawn at will.

(a) The most important variable is the temperature of the heater. The heater current is controlled by a rheostat and measured with a meter. The platinum loop is operated at a red glow temperature when small irregularities can be seen in the detail of the surface of the metal.

(b) The points in the cycle where heating current is stopped and the final strong pull commenced should be adjustable to fine limits, and in the instrument shown in *Plates 35.1* and *35.2* their positions can be noted on mm scales. The contacts controlling solenoid current and heater temperature should perform these functions through intermediate relays.

(c) The strength of the initial and final pull can be made adjustable. The initial pull should be rather weak, between 10–50 g, increasing to a maximum of over 1,700 g, according to Alexander and Nastuk¹⁵.

(3) When electrodes are drawn, the velocity and mass of the plunger combine to deliver a considerable blow to the instrument which may cause breakage of tips as a result of the vibration. This may be reduced by providing a friction break or cushion of soft plastic material. A new development of the apparatus by B. H. C. Matthews employs the solenoid to produce terminal deceleration as well as the tractive pull. This is done by reducing the length of the solenoid so that as the iron mass of the plunger passes through the direction of the field reverses. To arrest further movement a stepped ratchet brake is attached to the end of the plunger which

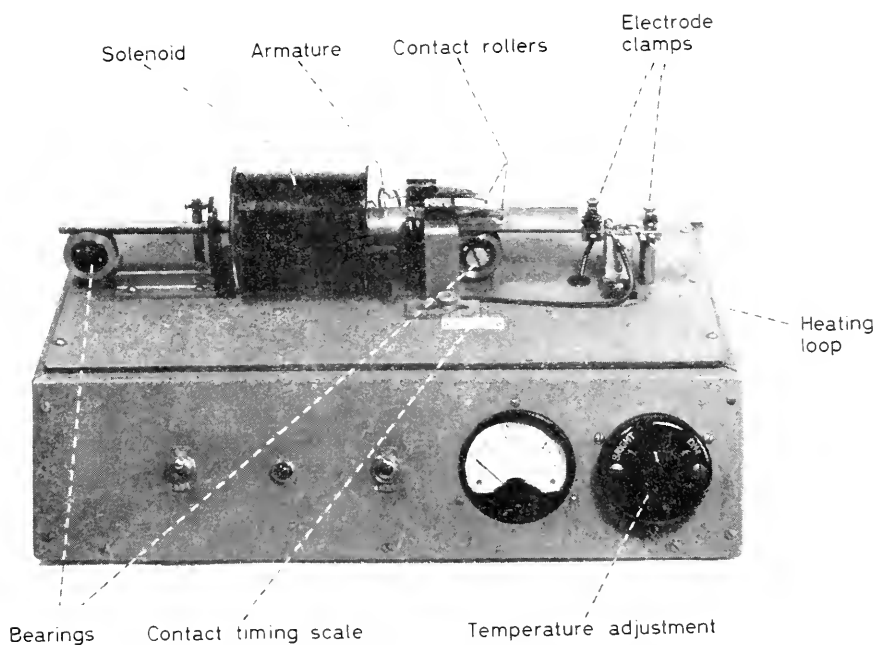


Plate 35.1 General view of microelectrode pulling machine constructed by R. H. Cook

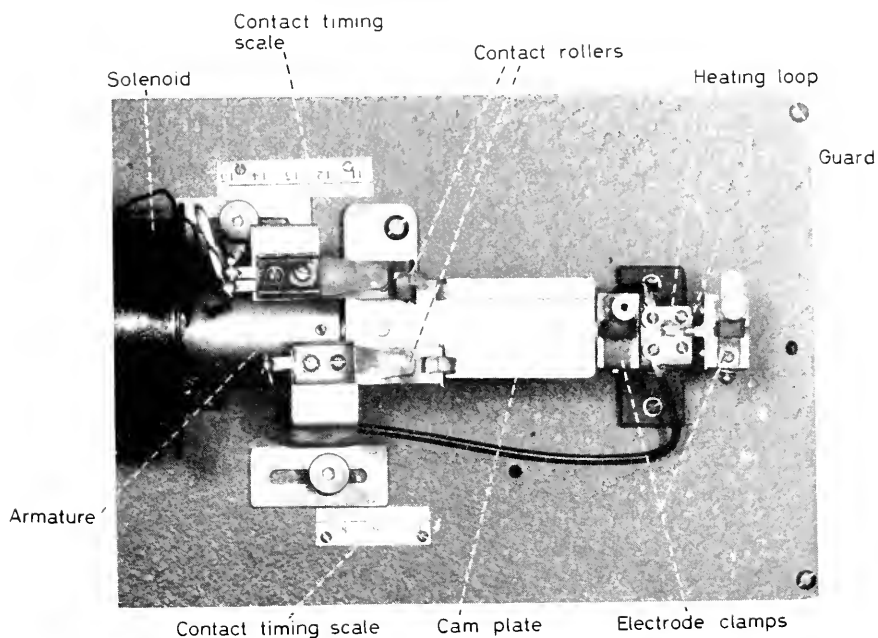


Plate 35.2 Detail plan view

(To face page 540)

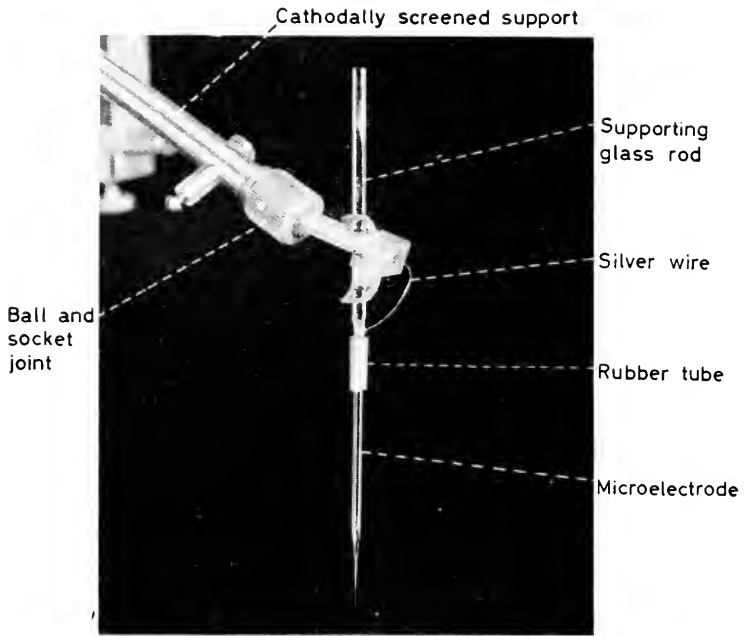


Plate 35.3 An adjustable mounting arrangement for microelectrodes

MICROCAPILLARY ELECTRODES FOR INTRACELLULAR RECORDING

engages a gravity arm. This arm is lifted over the teeth and also actuates a switch which cuts off power to the solenoid.

This apparatus is discussed in some detail as it is one to be recommended for the reliable production of a variety of microelectrodes.

Fine microelectrodes are usually made of Pyrex or Phoenix glass, the latter having a slightly lower softening point. For constant results tubing of fixed size and proportion should be used. The glass recommended by Alexander and Nastuk¹⁵ is 2 mm outside diameter and 1.3 mm internal diameter; but glass from 1.2–3 mm has been used. Glass of constant size can be ordered from manufacturers or hand-produced, cut into constant lengths of about 4 cm and selected to specified limits with a micrometer.

Examination and testing—The initial setting up of an electrode puller can be laborious, and systematic exploration is advisable. Electrodes drawn are first examined with a low power microscope when gross features are determined. The actual tip itself is often difficult to see clearly because of interference fringes which, at times, produce an apparently undulating surface or sharp taper.

It is advisable to fill electrodes as soon as possible. They are re-examined when full, with a water immersion objective, usually 1/6 in., when the image of the tip is clearer and the artefacts seen earlier are not apparent. The resistance of the electrodes is then determined, but whether the electrodes are in fact effective can only be decided by direct experiment.

Methods of filling capillary electrodes with electrolytes

It is advisable to fill electrodes with fluid soon after they are formed, as dust and moisture tend to adhere to them and even enter the lumen.

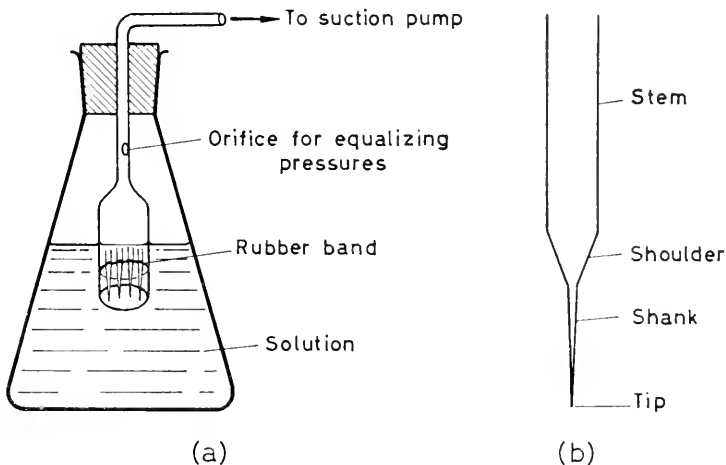


Figure 35.4 Arrangement for filling and storing microelectrodes

Electrodes are filled in the vertical position tip downwards. A simple method is shown in Figure 35.4 where the electrodes are secured by an elastic band around the lower end of a glass cylinder held by a rubber stopper in a large flask, which can be evacuated.

The earliest method of filling microelectrodes was to boil them for 30 minutes or more in 3 M potassium chloride; however, this may cause damage to the fine tips. A more efficient and less destructive process is to boil under the reduced pressure of a water pump. The solution is heated and the flask evacuated for 5–10 minutes. The bubbles are fine and the process takes less time than boiling at atmospheric pressure. Heating can be avoided simply by evacuating the flask containing the filling solution and electrodes, clamping the inlet and leaving for some hours.

Other slightly more elaborate methods have also been described, and may have to be used in some cases. In one of these (Tasaki, Polley and Orego¹⁶) the capillaries are first gently boiled in alcohol at reduced pressure. They are then placed in distilled water for a few minutes and finally transferred to 3 M potassium chloride which may also be evacuated. The method described by Caldwell and Downing¹⁷ is probably the most gentle of all. The electrodes are allowed to fill with distilled water by capillarity and the large air bubbles in the stem eventually dislodged with a probe. When filled they are transferred to, and left in, 3 M potassium chloride.

A method of producing pre-filled microelectrodes has been described by Kao¹³. Before drawing the electrodes the tubes are filled with 3 M potassium chloride and the ends left immersed in this solution during the drawing process: successfully filled electrodes result. A variation of this method is described by Crain¹⁴.

Once filled with electrolyte media, either physiological fluid or 3 M potassium chloride, the useful life of stored electrodes is about one week. Deposits and growths appear, and often the tips are found broken. The best way of preserving is possibly in boiled distilled water or alcohol. Among the methods which have been employed in attempts to prolong the useful life of microelectrodes are storage in the dark in a refrigerator, and an addition of small quantities of bacteriostatics or antiseptic such as methylene blue, which dissolves very slightly in 3 M potassium chloride. Their efficacy is doubtful. It is advisable to reduce the number of examinations and the transfer of electrodes to a minimum, since mechanical agitation, surface tension forces, and drying can cause damage.

Dimensions of electrodes required for intracellular observations

The size and shape of the microelectrodes used for intracellular work vary and a compromise is often reached between three factors; the nature of the tissue, the physical properties of the electrodes, and those electrodes which can be made.

The size of cells varies considerably with the corresponding variation in the size of acceptable electrodes. Crab muscle fibres continue to function with electrodes over 50 μ in diameter inserted, but at the other end of the scale small neurones 10 to 20 μ in diameter, such as are found in the lateral geniculate body, the spinal cord, and peripheral nerve, do not react well to penetration. It is possible that the leakage around an electrode is insignificant for large cells but may cause small cells to depolarize and lose much of their contents (Nastuk and Hodgkin⁷; Tasaki, Polley and Orego¹⁶).

To minimize damage, electrodes of less than 0.5 μ with a taper of 1/60 were used by Ling and Gerard³. This taper probably refers to the

shank of the electrodes. The majority of the figures quoted, however have shown a much sharper taper at the tip, and it seems that electrodes with gradients not greater than 1/10 are usable for most purposes (*Figure 35.5*): thus when such electrodes are introduced a distance of 10μ the

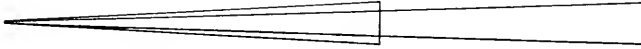


Figure 35.5 Illustration of tapers 1/10 and 1/5

actual puncture will be 1 to 1.5μ diameter. This does not seem to inconvenience some relatively small cells such as motor neurones which apparently tolerate the insertion of electrodes with tips of 1μ (Eccles⁶). Electrodes with gradients of 1/10 are more readily produced than those with lesser tapers, and the typical electrode shown by Alexander and Nastuk¹⁵ produced by their instrument is of this kind. The shorter tipped electrodes will have a lower electrical resistance, which will minimize errors in measuring both voltage levels and rapid signals.

The strength of the electrode may be an important factor in some kinds of experiment, such as the study of cells within ganglia (Eccles¹⁸). It is possible that the strength may be improved by the use of thick-walled tubing. When studying cells within the depths of a tissue it may be necessary to grade the shape of electrodes in several steps as, for example, by successive drawing under microscopic control (Brock, Coombs and Eccles¹²). For the purpose of electrical measurements errors will be minimized if the electrodes have as low a resistance as possible.

Methods of measuring the resistance of microelectrodes

The resistance of an electrode immersed in a conducting medium is readily determined by measuring the effect of a known resistance introduced into the circuit. Three such methods are shown in *Figure 35.6*. The experimental arrangement illustrated in (a) is suitable when isolated tissues are studied in a fluid bath which is insulated from earth. The deflection is noted when a voltage (≈ 10 mV) from a source of low output impedance is applied between the bath electrode and earth, and determined again with a known resistance R connected between the input lead and earth.

If V_0 is the initial deflection and V_1 the deflection with R in the circuit, the electrode resistance is given by

$$R_e = R \cdot \frac{(V_0 - V_1)}{V_1}$$

The approximate value of the electrode resistance can be determined by connecting a bank of resistances in series with a switch and finding that value which reduces the initial deflection to half. The resistances used can be 1, 2, 5, 10, 20, 50, 70, 100 M Ω .

In the method illustrated in *Figure 35.6b* a voltage is applied between earth and the input lead and the effect of placing a known resistance R in series with the input determined. In this case

$$R_e = R \cdot \frac{V_1}{V_0 - V_1}$$

GLASS MICROCAPILLARY ELECTRODES

Again a bank of resistances can be used to find that resistance which reduces a deflection to half its initial value.

A method used by Frank and Fuortes¹⁹, illustrated in *Figure 35.6c*, does not require an external source of voltage. The voltage used is that already present in the circuit, which may be due to such factors as the potential of

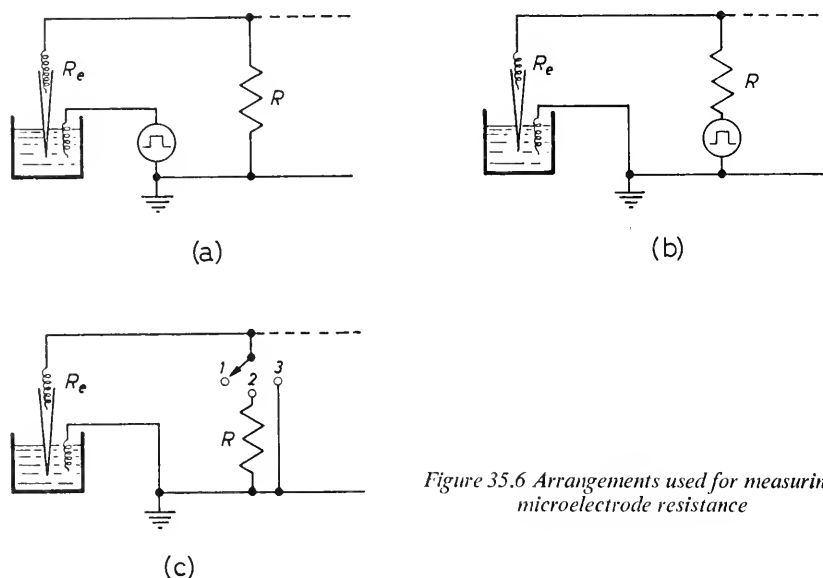


Figure 35.6 Arrangements used for measuring microelectrode resistance

the membrane electrode and junction potentials. A three-position switch is connected between the input and earth, in which the first position is open, the second connected through a known resistance R ($5\text{--}10\text{ M}\Omega$) to earth, and the third linked directly to earth. The input is connected to each position in turn and the potential differences noted. The electrode resistance is obtained either from the relationship

$$R_e = R \frac{(V_{1-3} - V_{2-3})}{V_{2-3}}$$

or

$$R_e = R \frac{(V_{1-2})}{(V_{1-3} - V_{1-2})}$$

Connection between electrode and amplifier

The simplest method of connecting the electrolyte within an electrode to the input of an amplifier is to insert a silver wire previously coated electrolytically with silver chloride. However, a preferable method is to connect the electrode to the metal junction through a Ringer-agar bridge: this will ensure a more stable electrode potential for the Ag-AgCl-electrolyte junction, as the electrode surface can be formed and left in place. The

electrode itself is inserted into a narrow rubber tube at the end of the Ringer agar bridge (*Plate 35.3*), which is filled with the Ringer's solution just before use. In the arrangement employed by Nastuk and Hodgkin⁷ the wire was held within the agar by means of a wax seal, which thus also reduced electrical leakage and prevented concentrated salt solution from escaping.

A more rigid form of holder made of plastic has been used, and consists of a split Perspex tube widening at its lower end, with a sliding ring to tighten the tube around the electrode.

The electrode should be rinsed with distilled water before use in order to remove external potassium chloride and increase the surface resistance.

Screening of electrodes

It has been found necessary to mount high resistance microelectrodes close to the amplifier input in order to improve the high-frequency response of the recording system. For this purpose, following the practice of Nastuk and Hodgkin⁷, electrodes are usually supported by the structure actually holding the input valve, thus reducing the distance between them to a matter of inches. The stray capacitance to earth of the wire connecting an electrode to the amplifier will tend to act as a high-frequency shunt. To reduce the stray capacitance still further the intervening link is shielded with a screen connected to the cathode of a 'cathode follower' input (*Chapter 18*). By this method the capacitance is effectively decreased, as only a small fraction of the input voltage appears between the grid and cathode, so reducing the flow of charge.

A cathodal screen will also serve to shield the input from interference. Electrodes of high resistance readily 'pick up' signals by capacitive coupling and a screen may be essential. Thus when stimulating and recording from the same muscle fibre Fatt and Katz²⁰ had to interpose a cathodal screen to shield the electrode almost to the level of the fluid contained in the tissue bath in order to reduce the recorded artefact.

A simple form of screen can be provided by a cathodally connected metal tube around the electrode. A convenient form of this is shown in *Figure 35.7a*. The electrode is attached to a glass tube containing Ringer-agar or a silver wire, while the tube is tightly held by a polythene support within a metal tube linked to the cathode. The cathodal screen is extended by an inner sliding tube which while allowing access for connection to be made can be brought down to shield the electrode.

Screening can also be provided by a metal film deposited on to the glass of the electrode surface (Marmont²¹) and a form of such an electrode was used by Frank and Fuortes¹⁹. Pt.Cl. or silver can be deposited on to electrodes before they are filled. Platinum was used by Marmont and a platinum resinate has been found satisfactory (Liquid Platinum, Johnson and Matthey F 104). When heated to 500°C a conducting film of platinum is deposited. The fluid is applied to the surface of clean and dry electrodes with a small brush or bristle, leaving a free margin of at least $\frac{1}{4}$ in. at the upper end of the stem and about 1 mm away from the tip. When the screen is to be cathodal a layer of insulating varnish must be applied to that part of the surface which will come into contact with tissue or bath fluid. Two

GLASS MICROCAPILLARY ELECTRODES

coats of 'Formvar' dissolved in chloroform can be used. A suitable form of electrode holder is shown in *Figure 35.7b*, which while supporting the microelectrode also connects the surface conducting layer to the cathode.

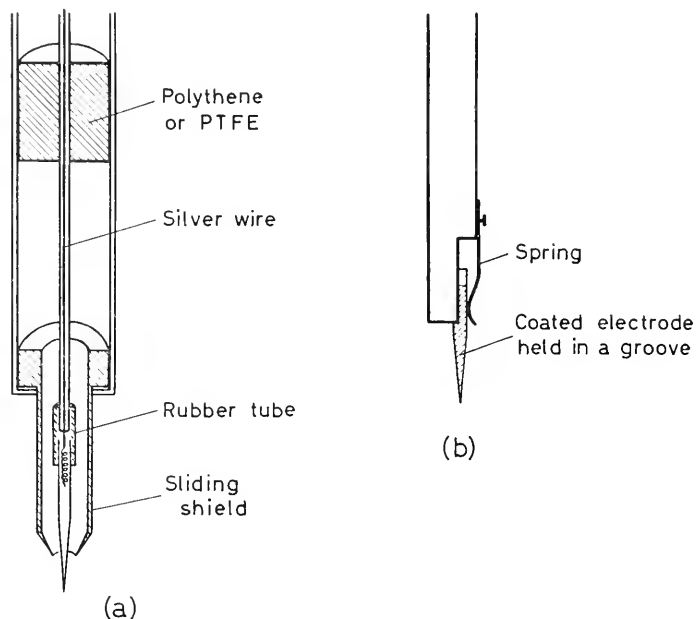


Figure 35.7 Screened microelectrode holders

Discussion on the process of drawing a capillary electrode

The process of drawing a capillary from a wider tube resembles that of the extension of a rubber tube; in both, as the length increases the diameter decreases. The elastic properties of glass, however, change greatly during a draw, as it passes from the fluid to the solid state. At high temperatures the elasticity, as measured by Young's modulus, is very low, and mounts rapidly on decreasing the temperature and approaching the solid state. The electrode is formed at a time when most of the physical properties of glass including viscosity, surface tension tensile strength and heat capacity are changing rapidly.

The 'shoulder' of an electrode (*Figure 35.4b*), is the point of most rapid taper, and the glass, because of the mass at this point, probably remains at the highest temperature throughout the process of pulling. The thin capillary drawn from the shoulder quickly loses heat, and in the short time of the pull it solidifies, as there is no evidence of sagging in most usable electrodes. The shank of the electrode is formed as the elasticity and viscosity of the glass is rising, so that the cooler the glass the more rapid will be the degree of taper obtained for the same rate of extension. Thus the cooler the glass, the shorter the length of the electrodes.

The ultimate tip of the electrode is probably formed from glass at the lowest temperature, near the transition point. The final tip will be formed when the force of extension exceeds the tensile strength of the terminal

fibre, and so will depend on the fibre diameter as well as the viscous and elastic properties at the prevalent temperature.

The last 10 μ of the electrode from the orifice are important to physiologists as it determines the area of membrane disturbed. This part of the electrode often shows great variability, usually an increased rate of taper, and may sometimes form a 'shoulder' just before the ultimate portion. Such a shoulder may be rather exaggerated when examined microscopically in air as a result of optical artefacts. The degree of taper at the terminal portion of the tip used by most workers appears to lie in the range between 1 : 8 to 1 : 10 (Nastuk and Hodgkin⁷; Alexander and Nastuk¹⁵; Tasaki, Polley and Orego¹⁶; Frank and Fuortes¹⁹; Fatt and Katz²²). This figure is derived chiefly from measurements of published photographs of micro-electrodes, and it is not known in fact whether greater variation is found in other satisfactory electrodes. Some figures which have been quoted showing a more gradual taper refer to portions of the electrode away from the tip.

The physical conditions under which the final submicroscopic tip is formed probably do not vary greatly when similar tubing is used. The shape of the tip will be determined by the initial temperature of the glass, the wall thickness and the velocity of pull. With thick walls the temperature to which the tube is heated will be higher, it will cool slowly and thus the electrode will be formed at a higher temperature. The velocity of pull will control the time allowed for cooling. As the tip is formed at a relatively low temperature the discontinuity in properties near the transition point may be responsible for the discontinuity of taper which is found in varying degree near the tip.

The actual proportions of wall thickness to lumen remain remarkably constant throughout the whole electrode. Thus the choice of the initial tube helps to determine the final dimension of the tip.

The mechanical strength of the tip may be important for some types of experiments. When greater strength is required the wall thickness may be increased by using thick-walled tubing. This is drawn at as high a temperature as possible which itself strengthens the glass. If the tip can be made without a terminal shoulder the electrode is probably stronger.

MEASUREMENTS WITH MICROCAPILLARY ELECTRODES

D.C. measurements

The physical size of a microelectrode tip gives it a high resistance which effects its functioning in several ways. A capillary tube with a lumen of 0.5 μ diameter will have a resistance of 5 M Ω for each μ of length when filled with Ringer's solution. Such electrode resistance will be placed in series with the resistance of the tissues and of the measuring instrument. When a voltage V is developed by the tissue, the value measured at d.c. will be determined by the proportion of the input resistance of the measuring instrument to the total resistance. Measurements of d.c. voltage within tissues are usually made in a bath of Ringer's fluid or *in situ*, when either the bath or the animal is earthed. The voltage can be measured with the aid of directly coupled differential amplifiers which possess good stability and an indicator.

GLASS MICROCAPILLARY ELECTRODES

The voltage recorded between the two input grids will be the sum of all the voltages in the external circuit, and besides the voltage due to the tissues present will also include the P.D. from metallic contacts of dissimilar metals at switches and junctions, electrode potentials of metals in aqueous solutions, and diffusion potentials in the fluids. When measuring the P.D. it may not be sufficient to determine the value of the potential deflection from the level achieved when both inputs are shorted with a wire to earth, unless care is exercised to reduce unwanted additional potentials. The sum of the total P.D. can be reduced both by the choice of materials forming junctions and by arranging that the unwanted potentials are placed symmetrically to oppose and cancel each other. Thus when a switch or other circuit element is introduced in one lead similar wires should be used to establish connection to both ends; and when a single junction is made, another identical one should be placed in the other lead. When a short circuit is to be established it is best placed as near to the point of measurement as possible so as to reduce the number of sites where potentials may be introduced.

Electrodes—An important source of e.m.f. in a circuit may be the electrode potentials at the metal fluid junctions. To reduce these, two similar junctions are arranged to oppose, with precaution to keep them stable. The junctions

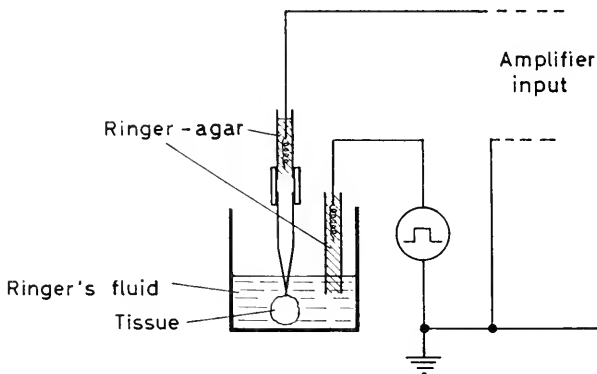


Figure 35.8 Arrangement for measuring with microelectrodes

commonly used for physiological experiments are of Ag-AgCl-Ringer. The Ringer can be made up with 2-3 per cent agar and pairs of electrodes kept together. One electrode is introduced into the fluid of the tissue bath and the other connected to the microelectrode filled with 3 M potassium chloride or Ringer's fluid. In this way fluid junction and electrode potentials are kept at minimal values (*Figure 35.8*).

Silver chloride working electrodes can be prepared in any size or shape required and are made from silver wire or platinum electrolytically coated with silver. Silver halide electrodes have been discussed recently by Janz and Taniguchi²³. The commonest form of Ag-AgCl electrode has a 'plum' or purple coloured surface and is relatively stable if carefully prepared with respect to current density and purity of solutions and metals. Silver is deposited on to platinum in a 1 per cent solution of potassium silver cyanide, and the electrode must be washed at least 12 hours in running water to

remove cyanide. The silver is chloridized at the anode, often in a bath containing isotonic sodium chloride or Ringer's fluid. However to obtain the most stable electrode a current density of approximately 5 mA/cm^2 of surface is passed in a solution of 0.1 M hydrochloric acid for 1–3 hours. By reversing the current several times the effective surface can be increased. Pairs of electrodes are short-circuited together and their tips immersed in Ringer's fluid. After some hours or days any initial difference of potential is eliminated. Silver chloride is soluble in concentrated potassium chloride so that for the most stringent condition the latter is not advisable as a direct connection. In order to reduce fluid junction potentials an intermediate potassium chloride bridge may be used if the electrode surface is placed in a weak chloride solution. Another possibility is the use of ammonium nitrate.

If direct sunlight be avoided the potential fluctuates less than 1 mV with variations in the ambient light. The potential will vary in the presence of some dissolved constituents such as chloride, bromide, iodide, or oxygen, when the electrode will tend to become more negative.

Fluid short-circuits—It may be necessary sometimes to use a fluid short-circuiting arrangement to examine the properties of certain electrodes or of artificial membranes. For this purpose a 3 M potassium chloride electrolyte bridge serves best, either connected between the fluid of the electrodes or between one electrode and the intermediate test or bath solution. It is not advisable to operate such a fluid switch by means of a glass tap, as the resistance of a greased tap with 3 M potassium chloride is often only a few megohms when closed. A convenient bridge to test electrode properties was used by del Castillo and Katz²⁴ in which a short arm filled with 3 M potassium chloride projected from the tube supporting the electrode, and could be connected to the bath fluid simply by lowering the electrode further.

Some common arrangements for measuring tissue voltages—It is often unnecessary to eliminate all sources of e.m.f. in a measuring circuit in order to measure d.c. voltages within tissues, as the voltages measured are essentially differences of potential between two points, such as outside and inside a cell or between a crushed and a functional portion of tissue. The amplifier can be brought back to its operating point by 'backing off', i.e. applying a voltage of the opposite sense in series by means of a calibrator. The most common cause of such potentials is that due to the use of dissimilar electrodes. Connection with a microelectrode is sometimes made with a chlorided silver wire inserted into its 3 M potassium chloride electrolyte, while the indifferent or bath electrode is formed by Ag-AgCl in Ringer's or tissue fluid. The resulting difference in electrode potentials has to be 'backed off'. While this can be readily done with a fluid bath some precautions are necessary when working with whole animals. A calibrator of low output resistance can be placed between earth and the indifferent electrode, usually a large silver plate, while care is taken to insulate the rest of the preparation from accidental earthing by insulating all clamps and supports.

An alternative point for 'backing off' is actually in the input lead. However this has disadvantages in that the whole apparatus must be well insulated, and can introduce distortion when recording rapid signals. A better method for re-balancing an amplifier is to include a voltage shift device

either in the cathode circuit of an input cathode follower or between the output of the latter and the input to the next stage. These methods are described in an earlier chapter.

Error in measurements of potential levels with microcapillary electrodes—It has been found that submicroscopic capillary electrodes may behave as if they were semi-permeable to certain ions; thus such electrodes sometimes give a potential which is a function both of the electrical potential present and of the concentration and type of ions at the tip. Observations of Nastuk²⁵, del Castillo and Katz²⁴ and Adrian²⁶ have shown that such electrodes filled with 3 M potassium chloride can commonly show potentials of 30 mV or more when immersed in Ringer's fluid, the inside being usually negative to the outside. Adrian²⁶ calculated that under these conditions the liquid junction potential at the tip between 3 M potassium chloride and Ringer's fluid should be only about 2 mV. He found that the potential depended on the concentration and type of ions in the external medium. The observations of del Castillo and Katz²⁴ and Adrian²⁶ have demonstrated that the potential at the tip is reduced or abolished when the resistance falls, consequent on breaking the tip, and, conversely, rises when there is an increase of resistance. The potential disappears when the composition of the external and internal medium is the same. This behaviour of electrodes has been ascribed by these authors to a blockage in the tip, possibly by protein, causing a reduction of the pore size. Electrodes in this condition may allow some ions to pass more readily than others, depending on the charge on the surface of the glass, and on the size and charge of the ions. Another mechanism which can be suggested is that the glass wall of the electrode behaves as an ion electrode for alkali ions similar to a pH electrode for hydrogen ions. Such an electrode will have a high resistance, which may be short-circuited by the electrolyte of the pore. When the resistance of the pore is increased the short-circuiting is less effective and the measured potential may increase. Electrodes which show a potential at the tip generally have a high resistance.

The result of this type of behaviour is that measurements of potential difference may show considerable error if the composition of the medium around the tip varies. The magnitude of the variation will depend on the actual electrode used, but with microelectrodes of resistance greater than 10 M Ω the resting membrane potential of muscle cells can be reduced by about 30 mV (Adrian²⁶) or occasionally increased. Sometimes electrodes may show such properties only intermittently, and del Castillo and Katz²⁴ describe changes in potential and resistance while actually inside a muscle fibre. They also found that electrodes with such potentials behaved as rectifiers, allowing current to pass more readily inwards than outwards.

To avoid errors due to such properties of the electrode tip it is necessary to measure the electrode resistance and potential regularly. Adrian used only electrodes with a potential of less than 5 mV but with resistance greater than 5 M Ω . These values can be determined during an experiment, when an increase in resistance may be indicated by an increased noise level and greater susceptibility to electrical interference.

The method employed for filling and storing electrodes may be a factor inducing these properties. Nastuk²⁵ found that boiling in 3 M potassium chloride caused electrodes to develop such potentials, this may even happen

MEASUREMENTS WITH MICROCAPILLARY ELECTRODES

as a result of storage. In concentrated potassium solution some of the ions may perhaps enter the structure of the glass and it is possible that this may be linked with the actual mechanism of the potential. The electrode properties of glass alter with constitution and increase to some degree with respect to the ions contained. Entry of ions also makes glass more brittle and this may be one cause of the short life of microelectrodes. Nastuk filled electrodes by means of an alcohol replacement method, and long-term storage in distilled water or alcohol may be advantageous.

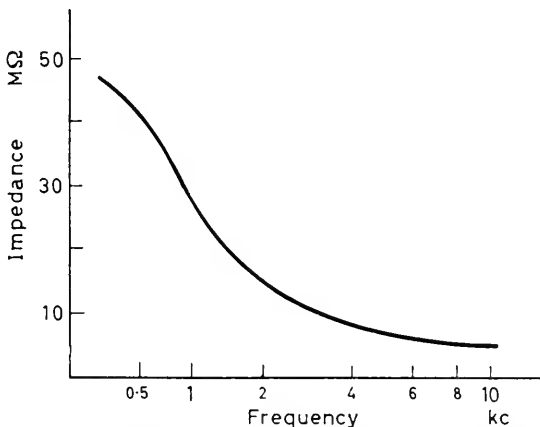


Figure 35.9 Plot of impedance of a microcapillary electrode against frequency: electrode filled with 3M KCl. Drawn from data given by Tasaki²⁷

Measurements of rapid signals—The impedance of microcapillary electrodes is frequency sensitive. Figure 35.9 shows the type of curve obtained when the impedance is measured with the electrode tip immersed in Ringer's fluid (Tasaki²⁷). As the frequency is increased to 5 kc impedance falls to approximately one tenth or less of that at d.c. The cause of this is that at higher frequencies the microcapillary no longer behaves as a simple pore electrode. The impedance of the glass wall falls and becomes comparable to that of the electrolyte filled lumen. The wall of an electrode suitable for intracellular puncture may be only 50 $\mu\mu$ thick at the tip. It behaves as the dielectric of a capacitor with concentric plates. A capacitance of this order however extends up the shank of the electrode, because as the effective separation of the plates increases with greater wall thickness the total area per unit length also increases with increasing diameter. These properties of microelectrodes were described and examined by Nastuk and Hodgkin⁷.

When a microelectrode is inserted into a cell within a bath of fluid or a mass of tissue, the source of voltage—at the membrane—is connected between the fluid at earth potential and the tip of the electrode. The resistance of the electrolyte in the capillary, R , is connected in series with the source, which it links to the input of the amplifier, while the capacitance of the wall of the electrode, C , is placed in parallel with the source, between the lumen and the earthed fluid (Figure 35.10a). This can be represented by the arrangement in *b*.

If a voltage step, V_i (Figure 35.10b), is applied between the input and earth, the output voltage, v_o , across the capacitor will not attain a steady value

GLASS MICROCAPILLARY ELECTRODES

until it is charged, so that the output voltage will follow an exponential course, The current, i , at any instant, t , flowing through a resistance, R , causes a voltage drop iR

and
$$iR = V_i - v_o \quad \dots (1)$$

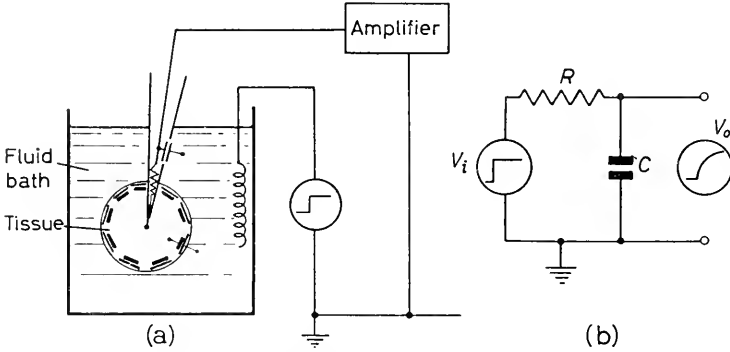


Figure 35.10 (a) Representation of a recording arrangement;
(b) equivalent circuit of (a)

This current charges the capacitor C

$$q = Cv$$

$$\frac{dq}{dt} = C \cdot \frac{dv}{dt} \quad \dots (2)$$

but
$$\frac{dq}{dt} = i$$

Substituting for i from expression 1, expression 2 becomes

$$V_i = v_o + RC \cdot \frac{dv}{dt} \quad \dots (3)$$

The voltage v_o at any time t is less than V_i by an amount equal to a constant of the circuit (RC) times the rate of change of voltage at that time (see also Figure 35.12 and Chapter 3). When the voltage reaches a constant level $dv/dt = 0$ and $V_i = v_o$,

The solution of expression 3 is

$$v_o = V_i(1 - e^{-\frac{1}{CR} \cdot t}) \quad \dots (4)$$

See Figure 3.9

When $t = CR$ seconds

$$v_o = V_i \left(1 - \frac{1}{e}\right)$$

i.e. at $t = CR$ the voltage has risen to 63 per cent of its final value. The product RC , the time constant, can be determined from the response curve on applying voltage step and finding the time taken to reach 63 per cent of the final steady value.

MEASUREMENTS WITH MICROCAPILLARY ELECTRODES

Thus when a brief signal is applied to the tip of an electrode the full voltage may not be recorded at the output unless the capacitance has had time to be charged. The time constant is a measure of the speed of the response, and can be used to construct the frequency response curve to steady-state signals. At any frequency the fraction of the voltage measured will be determined by the proportion of the reactance of the capacitor to the total impedance (*Figure 3.24*),

$$\text{i.e.} \quad v_o = V_i \cdot (-j) \frac{X_c}{Z}$$

where X_c is capacitive reactance and Z is total impedance.

As described in Chapter 3, the frequency at which the response is reduced by 3 dB's or to 71 per cent is obtained from

$$f = \frac{1}{2\pi \cdot C \cdot R}$$

where f is the required frequency in c/s.

It can be seen by substitution that when the time constant is 50 μsec the response will be reduced to 71 per cent at 3.2 kc, and with 100 μsec at 1.6 kc. In order to reduce the shunt capacitance of the recording system the method of cathodal rather than earthed screening can be used.

To examine the properties of an electrode under conditions similar to those when recording from within a cell, Nastuk and Hodgkin⁷ devised an arrangement which simulates this closely. A potential is applied to the tip of an electrode while a length of shank nearby is surrounded by fluid at earth potential. An electrode is passed through a globule of saline supported on a wire ring connected to earth. The tip just touches the surface of the fluid in a bath beneath, and the voltage source is connected between earth and the bath. The capacitance of the electrode wall in the fluid was about 1 pF per mm of length. The capacitance of the shank in air is a small fraction of the value in saline, as effectively the wall capacitance is in series with the air capacitance to earth. However with cathodal screening the total capacitance usually ranges between 3–10 pF. Of this total 1–3 pF may be due to the cathode-follower input stage while the rest is principally that of the electrode.

Microelectrodes used in a bath of tissue fluid with cathodal screening generally have a time constant of 30–100 μsec . The degree of distortion introduced depends on the activity studied and will vary with temperature and other factors. For example, published data for the action potential of frogs muscle show that the height may be reduced approximately 1–8 per cent under these conditions. At low temperatures greater time constants can be used. The greatest effects of time constant distortion are on the rate of rise of a potential. Generally a time constant of 30 μsec will introduce little distortion in recording frog muscle potentials. It has been estimated that more rapid events, such as the action potential of some mammalian neurones—0.5 μsec duration, will require a time constant of 15 μsec to avoid appreciable distortion (Woodbury²⁸)

The time constant can be reduced by decrease in both the input resistance and capacitance. It is difficult to decrease the resistance without increasing the dimensions of the electrode. Very small tips may be required for work

GLASS MICROCAPILLARY ELECTRODES

with small cells, such as internuncial neurones, and possibly electrodes with a sharp taper near the tip may be of help. However, Frank and Fuortes¹⁹ found it necessary to use electrodes with resistances 50–100 MΩ or more. To reduce the time constant, electronic correction can be used, which can give an effective input capacitance less than 1 pF.

When a voltage step is applied to a fluid bath with the microelectrode immersed, as in the recording position, the voltage waveform shows an initial sudden step followed by an exponential rising portion up to the final level (Figure 35.11). The initial step is caused by the displacement of charge

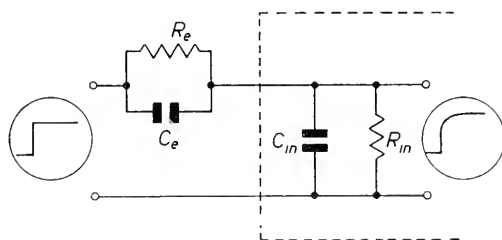


Figure 35.11 Equivalent circuit of a microelectrode dipped in a fluid bath through which it receives a pulse

in the wall capacitance in series with the voltage source. The height of the step is determined by the ratio of the electrode wall capacitance (C_e) to the input capacitance of the cathode follower (C_{in}), and is proportional to $C_e/(C_{in} + C_e)$. The slowly rising portion represents the process of charging both capacitors through the electrode and input resistances in parallel, with a time constant therefore of

$$(C_{in} + C_e) \left(\frac{R_{in}R_e}{R_{in} + R_e} \right)$$

where R_{in} = input resistance of amplifier and R_e = electrode resistance. If the input resistance is much larger than that of the electrode then this simplifies to:

$$(C_{in} + C_e)(R_e)$$

The time constant can be determined from this.

Reduction and correction of distortion—It has been shown that the output differs from the input voltage in expression 3 by an amount

$$RC \cdot \frac{dv}{dt}$$

All three factors can be used to improve the response of the system.

A correction based on the rate of change of voltage has been applied by Woodbury²⁹ at the output of the amplifier. The voltage from this is differentiated by means of a resistance-capacitance network of short-time constant, whose output is then added to the original in a mixer stage.

A method of reducing the effective capacitance has been described by Solms, Nastuk and Alexander³⁰ and similar methods have been developed by McNichol and Wagner³¹ and Haapanen and Ottoson³². Positive feedback is applied from the output of the cathode follower back to the input so that

MEASUREMENTS WITH MICROCAPILLARY ELECTRODES

the response to a square pulse is improved. Effective time constants under 10 μ sec can be obtained.

The most common method of minimizing distortion is to use electrodes of the lowest resistance possible with the provision of cathodal screening.

Graphical and mechanical methods of correcting for time constant distortion—In order to examine and correct for distortion introduced by the response time constant of the input the method described by Burch³³ and by Lucas³⁴ or the mechanism given by Rushton³⁵ can be used. The method is based on the fact that the subtangent of an exponential curve is constant

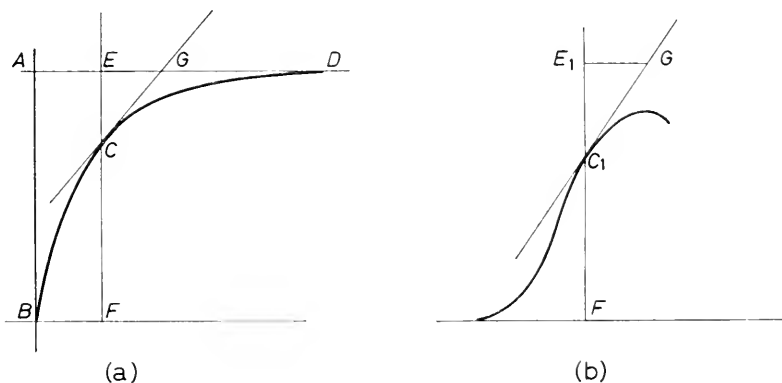


Figure 35.12 Graphical correction for time constant distortion

at all parts of the curve: if the value of the subtangent be found for an experimental arrangement, it can then be used to correct for distortion in a recording.

In Figure 35.12a the response to a voltage step is given by the exponential curve BCD , and the level which it will finally attain is AD . At any point C , the value CF is less than the full response FE by an amount CE . If the tangent at C be drawn to cut AD , then EG is the subtangent. The value of EG (in sec) is constant at all points of the curve. When EG is known, CE can be determined.

A correction can now be applied to an experimental curve (as in Figure 35.12b) obtained under the same recording conditions of resistance and capacitance. At any point C_1 , a vertical line and a tangent are drawn. The point E_1 is marked on the vertical where the horizontal line E_1G is equal to the subtangent EG . The point E_1 represents the true voltage at this time on the response curve. The correction is performed at successive points on the curve.

Recording within a mass of active tissue—Records may be required from a structure within a tissue mass, when activity is widespread and varies from point to point. The electrode may have to pass through one active region in order to study another beyond it. It is possible that activity may be recorded by capacitive coupling through the walls of the shank of the electrode (Hodgkin⁹). The degree of interference will depend on the relative voltages of the source and interfering region, and on the dimensions of the latter.

A simple experiment to illustrate this phenomenon can be performed by applying a 10 kc sine wave to a bath of fluid while an electrode is introduced:

GLASS MICROCAPILLARY ELECTRODES

the voltage recorded at 5 mm depth may be double that at the surface. However in order to examine the behaviour of an electrode under conditions resembling those when recording, a method similar to that of Nastuk and Hodgkin⁷ can be used. This was done by passing an electrode through a piece of agar-Ringer 3 mm thick, placed on a silver plate with a hole in the centre. The tip of the electrode was immersed in Ringer's fluid below the silver plate (*Figure 35.13a*). The potentials were compared when the tip was given a voltage step with the agar earthed, and when the agar received the pulse with the tip earthed. (The responses were as sketched in *Figure*

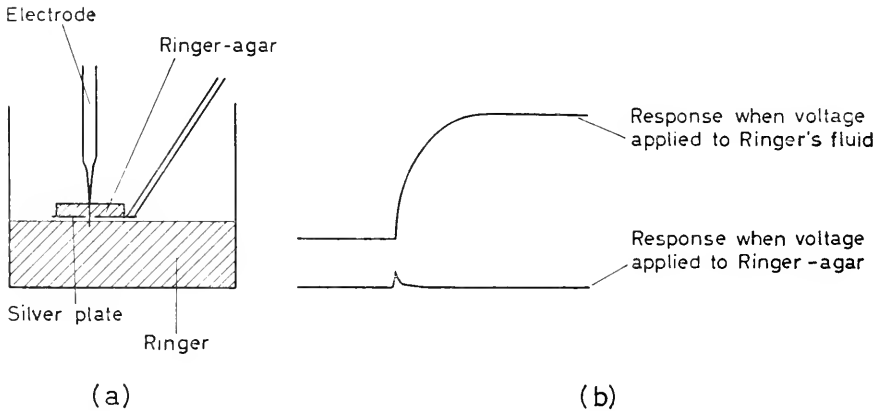


Figure 35.13 Text for the recording of a signal through the wall capacitance

35.13b; of exponential form and of similar time constant. The ratio of their heights was approximately 20 : 1. The input capacitance of the cathode follower was 3 pF under these conditions, but if this were smaller, a greater potential would be recorded through the wall of the electrode. It is not likely that this kind of interference would be significant when recording large intracellular potentials, but if small extracellular voltages are examined and a large region along the shank were to become active it is possible for some significant interference to occur. This type of interaction may be more serious with other types of electrode, such as varnished metal needles, which may possess thinner insulation and greater wall capacitance.

When it is necessary to reduce interference of such kind electrode screening can be used. A metal such as platinum or silver can be deposited on to the electrode down to a millimetre or less from the tip, with an additional layer of insulating varnish: the method has been described earlier. It will usually be necessary to use cathodal rather than earthed screening to minimize capacitive distortion.

MICROCAPILLARY ELECTRODES FOR EXTRACELLULAR RECORDING

Saline conducting medium

The voltages recorded in the conducting medium surrounding active tissues are but a fraction of those generated across the tissue cell walls themselves.

The actual values will be determined by the proportion of the external resistance to the total resistance of the circuit, so that small masses of tissue tend to show greater interstitial voltages than larger ones. Values of potential difference so measured within the central nervous system are generally less than 1–2 mV, though larger potentials have been recorded.

The most common arrangement used for recording is to measure the P.D. between an electrode in the region of activity and an indifferent electrode placed on inactive tissue, which can also be earthed: at the inactive point the tissues may be merely quiescent or deliberately destroyed. The difference in electric potential between the electrodes at any instant is that due to the flow of current through the medium between them: its value is given by the integral of the P.D. along a line between the two points. It can be regarded as the sum of all the products IR taken over small finite distances along that line. The studies of Lorente de No³⁶ have shown that the potential in a volume of conducting fluid approximates to the second derivative of the tissue membrane potential and is due to the transverse membrane current.

The smallest kind of microelectrode is not often used for extracellular studies. The potential recorded from such electrodes tends to be dominated by the activity of a single cell nearby, probably because it can approach closely without distortion of tissues. However electrodes of 1–3 μ can be used to examine the external field of individual cells.

Resistance of electrodes

When recording extracellularly the electrical noise of the electrode may be comparable in voltage to that of the signal, giving a relatively low signal-to-noise ratio. The noise due to the resistance R of an electrode is given by $1.8 \times 10^{-8} (R)^{\frac{1}{2}}$ volts R.M.S. for the audible range of frequencies. In order to reduce the noise level the resistance should be kept at a minimum. When electrodes of about 2 μ diameter are employed they can be filled with concentrated sodium chloride solution (5–6 M). This is used in preference to potassium chloride as the effect of any leakage from the tip will be less harmful. Electrodes of this kind have resistances of the order of 1 M Ω with a R.M.S. noise level of approximately 20 μ V.

When electrodes larger than approximately 3 microns tip diameter are employed they are usually filled with isotonic solutions, either Ringer's or sodium chloride, so that leakage of contents will not be deleterious. The diameter of the tip is made as large as can be tolerated in order to reduce the resistance. In addition a silver wire can be introduced as far down as possible into the shank. Some further details of this are given later.

Reduction of the electrode resistance will not effect 'noise' generated by tissues. Such noise is recorded in a mass of tissue, as the central nervous system, particularly when the electrode is close to a cell. Under some conditions the tissues themselves may have a high resistance which may generate noise in the circuit.

Some methods of making microelectrodes for extracellular use

Glass microcapillary electrodes for extracellular use can be made by any of the methods described for the smaller type. Mechanical methods however are much more satisfactory than hand drawing.

This larger kind of electrode is also easily made from the smaller ones simply by breaking the tip. Controlled damage can be produced by carefully lowering an electrode on to a solid surface with the aid of a micromanipulator. The electrode can also be stabbed several times through cotton wool, filter paper or a jelly containing a suspension of fine carborundum powder.

Another method of breaking the tip is by passing a high-voltage arc discharge between it and a carbon or metal electrode brought close, and this can be done when the microelectrode is filled with conducting fluid. A motor car induction coil may be used. However the method is not a good one as it may cause cracks along the shank.

A simple and reliable method which has been used to enlarge an electrode tip is to expose it to fumes of hydrofluoric acid in the neck of a bottle containing a 48 per cent solution. A brief exposure of one or two seconds at a time is made, followed by washing in water (Tobias and Bryant³⁷). Electrode tips can also be ground on a rotating fine abrasive wheel of India or Arkansas stone. The electrode is approached with a micromanipulator at a slight angle, pointing in the direction of rotation. The process can be observed with a dissecting microscope and the point where it makes contact is immediately seen, as the electrode moves laterally. The sound of the abrasive action can also be used as a guide.

Glass microcapillary electrodes with saline and metallic conductors

The most common way of reducing the resistance of saline-filled microcapillary electrodes is to introduce a silver wire as far as possible: if the wire protrudes at the tip both can be cut off. However it is not possible to insert a wire into very fine capillaries, so that electrodes of this type are usually greater than 20 μ diameter. With electrodes too small for the wire to reach the tip, the major portion of the resistance remains.

Electrodes with metallic conductors as far as the tip, while suitable for recording fluctuating signals, are usually not reliable for measuring steady potential levels as the electrode surface is small and may be subjected to changes in composition of the medium, and to chemical action: this will cause potential drift. Such difficulties were overcome by Hodgkin and Katz³⁸ by inserting a bright silver wire into the saline-filled electrode from the tip to the fluid in the stem. This wire did not make direct contact with the grid wire. For rapid signals the impedance of the silver wire was less than that of the saline bathing it, which was thus effectively short-circuited. At d.c. it behaved simply as a saline-filled capillary electrode.

Tomita and Funaiishi³⁹ have described a form of saline-filled capillary electrode with a silver wire as far as the tip. They have found that the noise of the electrode was greater when it was not filled with saline. It is possible that the saline surrounding the wire near the tip served to increase the effective area of the electrode surface; alternatively the saline fluid may have increased the value of the parallel wall capacitance and caused a reduction in noise by reducing the high-frequency response.

Glass microcapillary electrodes with metallic conducting media

Electrodes of glass can be formed with conductors of various metals. Their value is found in the relatively low noise level and in an improved

high frequency response when compared with electrolyte-filled microelectrodes. Their properties at zero frequency however are not as stable as the electrode surface is actually in contact with the tissues.

The resistance of the metal conductor of such electrodes is small, but that of the electrolyte surrounding the tip is appreciable in comparison. The resistance of a $2\ \mu$ diameter wire of the commonly used metals such as silver, copper, tungsten, platinum, iron, tin and lead lies approximately between 5 and $70\ \Omega$ per mm. The resistivity of Ringer's fluids is of the order $10^2\ \Omega\ \text{cm}$,

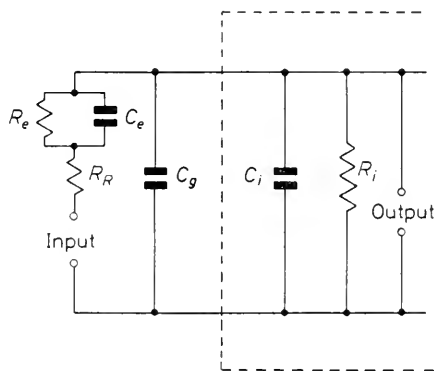


Figure 35.14 Circuit of a metal-filled microelectrode with tip immersed in a saline medium

R_e	electrode resistance
R_R	resistance of Ringer's fluid at tip
R_i	amplifier input resistance
C_e	capacitance of electrode surface
C_g	capacitance to earth of electrode insulation
C_i	amplifier input capacitance

but the dimension of the conducting path at the electrode approaches the size of the tip itself, and may be greater than $10^5\ \Omega$ for a tip $2\ \mu$ diameter (Svaetichin⁴⁰). The impedance of the electrode surface itself is frequency dependent, decreasing as the frequency increases. Electrodes of this kind generally behave as if they possessed a capacitance in series with the input (Figure 35.14).

Polarization—Metal-filled electrodes show polarization in varying degree when current is passed through them (Chapter 34). A polarizable electrode changes its potential with respect to the fluid medium in a manner not accounted for by Ohm's law when current passes, while a non-polarizable one does not do so. The polarization can be determined by measuring the potential difference between the electrode in question and a third indifferent electrode (Rothchild⁴¹).

Polarizable electrodes behave as if they had a capacitance in series with the input, and this capacitance may seem to be large. The tip of a $30\ \mu$ platinized platinum wire may appear to have a capacitance of $1\ \mu\text{F}$ when a direct current of $10^{-8}\ \text{A}$ is suddenly applied through a high resistance. This is due to the formation of an electrical double-layer at the surface and to electrochemical mechanisms which take some time to reach a steady state. The double-layer can have a capacitance of approximately $20\text{--}40\ \mu\text{F}/\text{cm}^2$

(Philpot⁴², Fricke⁴³ and Grahame⁴⁴). The electrochemical polarization capacitance is not constant and decreases as the frequency increases. This is probably because time is required for the electrochemical action to take place.

Some methods of constructing metal-filled electrodes

To construct a metal-filled electrode the metal may be introduced after the capillary is formed or the electrode drawn from pre-filled tubing. An example of the first method is that described by Weale⁴⁵. After the capillary is drawn a stout silver wire is jammed into the stem. The electrode is filled with aqueous silver nitrate solution and a chain of silver crystals deposited in the shank up to the tip, by means of electro-deposition. The process can be observed under a microscope. Such electrodes should possess non-polarizable surfaces in physiological fluids however their resistance rises in use, possibly due to the deposition of a resistant layer (Svaetichin⁴⁰).

The method of drawing a metal-filled glass capillary was described by Taylor⁴⁶. The metal and glass used are such that the softening point of the glass lies between the melting and boiling points of the metal. A simple combination is that of soft solder and soda glass. A piece of metal is placed in the bottom of a tube closed at one end, and heated. When the metal melts an oxide film forms which it is advisable to remove by pinching (with pliers) the bottom of the glass tube repeatedly—when the film will stick to the glass—until most of the oxygen at the bottom of the tube is used up. Capillary tubes can now be drawn in one or more stages containing metal with a bright surface, which should be inspected for cracks.

Electrodes with a silver alloy conductor can be made and those described by Svaetichin⁴⁰ have been used to study single cells. Their properties have been examined by Gray and Svaetichin⁴⁷. Tips of under $1\ \mu$ diameter can be formed under microscopic control from capillaries of approximately $20\ \mu$; the alloy is then electrolytically coated with rhodium to form a hemispherical cap, and finally platinum black is deposited. The latter forms a spongy layer with a relatively large effective surface which can be cleaned in acid when necessary, and reformed.

Electrodes such as these have a low noise level which was measured in many electrodes by Svaetichin⁴⁰, e.g. a tip of $5\ \mu$ diameter had a peak-to-peak (sic) noise of $10\ \mu\text{V}$. The impedance of these electrodes measured at 1 kc is approximately $0.2\text{--}1\ \text{M}\Omega$ for tips 3 to $10\ \mu$. Their high frequency response is found to be very good. However the potential level of the electrode is not stable. Gray and Svaetichin⁴⁷ found large drifts on first introducing the electrode into an electrolyte solution, and after 15 minutes the drift declined to approximately 1 mV per minute. These electrodes have also been employed for intracellular recording. A comparison of the membrane potential recorded with these and with saline-filled electrodes would be of value.

Another form of electrode has been described by Dowben and Rose⁴⁸ whose preparation is relatively simple and certain. It embodies a novel construction in that an alloy of indium—a glass-wetting metal of low melting point—is used. A microcapillary is drawn from a tube containing some indium, and when gently heated the metal can be pushed down to fill the electrode. The indium alloy at the tip is electroplated first with gold

GROUPS OF GLASS MICROELECTRODES

then with platinum. The properties of these electrodes are those of a platinum surface electrode.

GROUPS OF GLASS MICROELECTRODES

The grouping of microelectrodes may be conveniently considered to be of two types; two or more electrodes at a single site or a cluster of individual electrodes at a fixed distance apart. The first of these really consists of a single structure with a number of channels. Two channel electrodes were described by Renshaw, Forbes and Morrison⁴⁹ who employed them to study potential gradients in the brain. Structures such as this have been used to inject substances in the central nervous system (Kennard⁵⁰), to inject and record (Coombs, Eccles and Fatt⁵¹) and for surface application to muscle cells (del Castillo and Katz⁵²).

Construction of a two-channel capillary

Two glass tubes 5–10 mm outside diameter and 1 ft. long are bound together tightly with wire, rubber bands, or a cement (Picene). At one end both tubes are then fused and closed. Heat is applied a few in. from the

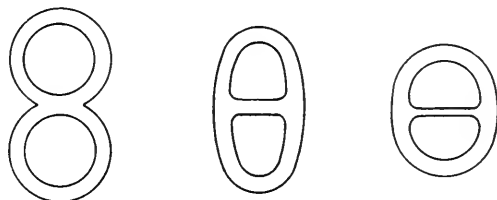


Figure 35.15 Stages in the lateral fusion of two glass tubes

end and when soft both tubes are blown so that they stick together. By repeated heating and blowing the tubes at this point can be fused together by varying amounts. The least degree of fusion is that when the tubes simply form a figure 8, while complete fusion gives a structure with oval cross-section containing two *D* shaped compartments back to back (Figure 35.15). As the heating proceeds the glass tends to flow from the dividing wall to the exterior, and if excessive may lead to a very thin and fragile partition. When the required degree of fusion is achieved the glass is drawn down to a diameter of 1–2 mm, while it is gently blown by mouth. Henceforth the tube can be treated as a single channel microcapillary and drawn down further manually or by machine. Often however, as a result of unequal stresses, the tube tends to flatten and occlude, and this can be prevented by applying a slight air pressure (4 in. water) during the draw. Another method of making a two-channel tube is to fuse a glass partition (e.g. microscope slide) across the middle of a wide-bore tube.

Electrical leakage—To prevent electrical leakage between the two channels at the stem an angled tube can be cemented into each channel by means of a low melting point wax such as Picene or dental wax. Alternatively small portions of the unfused parent tubes can be left and slightly separated⁵¹.

Leakage between the two compartments in the shank or near the tip may also cause difficulty. Leaks can be detected sometimes by comparing

the resistance between the two channels with the tip in saline and in oil. To some extent leaks of this kind can be avoided by using tubing with thick intermediate partitions. An added glass partition was found necessary by del Castillo and Katz⁵².

Some methods of modifying microcapillaries

Occasionally specially shaped tips may be required in one or more channel microcapillaries (Kennard⁵⁰). A tube may have to be broken or the intermediate wall removed a short distance, and it may be necessary to close an orifice partially. The ease of such operations will depend on the size of the tip.

The simplest method of breaking off pieces from the tip is by means of a fine steel microneedle carried in a micromanipulator: however, a more readily controlled method is a vibrating needle. This is mounted on the diaphragm of an earphone—the best being the balanced reed type—or directly on the electromagnet of an earphone. It is of advantage to have the needle vibrating in a plane rather than a cone. The microcapillary is supported as close to the tip as possible and the vibrator is carried on a micromanipulator. Power is supplied from an oscillator and the best frequency found by trial. The orifice can be closed by applying heat with a loop of fine platinum wire under microscopic control. For this purpose the operation can be carried out with the capillary mounted vertically and viewed end on through transmitted light. The contrast—and therefore the image—is improved by surrounding the capillary stem with plasticine.

Construction of capillaries with three or more channels

Glass capillaries with many channels can be made (Elson⁵³, Vis⁵⁴). Those with 3, 7 and 12 tubes are most readily drawn as these numbers of tubes form stable arrangements: the principle of the method is similar to that used for two-channel capillaries. Care must be taken to immobilize the tubes, and they can be pushed into tight fitting rings for this purpose. Heat should be applied gradually and kept to a minimum in order to avoid sharp gradients of temperature. To blow numbers of tubes in an array it is necessary to connect each tube individually by means of soft rubber tubing and a multi-way connection. To avoid lateral hindrance each tube is first drawn down to about half its diameter for approximately 1 in.

A capillary formed of three tubes has an additional channel in the centre, and the larger groups have many more channels than parent tubes though some may be severely distorted. Capillaries of this kind are fragile and not readily worked.

Fused quartz tubing with several compartments is manufactured commercially. It is used chiefly for purposes of insulation.

Multiple-bore tubes can be used for spacing arrays of electrodes and for extensive study of a small region of active tissues.

Clusters of capillary electrodes

Groups of electrodes may be required to measure the potential at points a short distance apart, arranged to cover an area or in some definite array. If such an area be at the surface of a tissue several micromanipulators can

GROUPS OF GLASS MICROELECTRODES

be used: however, if the study is to be made at a depth within it is usually necessary to arrange the electrodes with their longitudinal axes parallel so that they may be introduced in this direction. Tomita and Torihaina⁵⁵ have described an arrangement for two electrodes in which one is actually introduced through the other and both can be moved independently along their common axis.

A group of electrodes can be placed close together and parallel by using their stems to align them side by side. For this purpose the stem diameter must be made equal to the spacing required, and the shanks and tips have



Figure 35.16 Schematic representation of the alignment of electrodes

to be coaxial with the stem. Such electrodes can only be prepared by mechanical means. Inter-electrode spacing larger than the diameter is accomplished by interposing a spacer such as paper, foil or cover slip. With small spacing it becomes more difficult to make glass electrode assemblies with parallel stemmed electrodes less than 0.1 mm diameter: metal electrodes made of uniform wire may be of use.

Groups of electrodes with very small spacing can however be arranged by manipulating electrodes or their tips into place. For this purpose capillary electrodes are drawn by hand with a range of tapers and lengths, and the shanks of those suitable are cut off at the shoulder of the stem. The open end forms a small funnel which facilitates the introduction of wire connections.

Longitudinal spacing—When a series of collinear points at different depths are to be studied it is sometimes possible to introduce an electrode assembly along the line. In this case the electrodes can be held together with their tips placed at the required intervals along their shanks (*Figure 35.16a*). This form of spacing is simpler to arrange than that side by side (*Figure 35.16b*). One electrode is cemented lightly with wax across a thick hypodermic needle (*Figure 35.19*) with the aid of a small soldering iron. Some Vaseline is applied to a small area of another electrode a short distance from the thick end and the second electrode is placed along the first one. The second electrode is manipulated under a microscope until the tip is at the required position, the Vaseline being sufficient to support it in place while they are cemented together. The cement used should be very thin and spread by surface tension, but should not dry too rapidly. Durofix diluted with ethyl alcohol has been found satisfactory but it tends to cover electrode tips. A weak solution of gelatin can be used and is much less likely to block the orifice; it can be dried quickly under a lamp and eventually hardened if necessary in formalin. Further electrodes can be added in this manner. Groups of seven electrodes at a spacing of 100–150 μ have been constructed which are not much larger near the tip than some single electrodes.

GLASS MICROCAPILLARY ELECTRODES

Additional electrodes have to be spaced by reference to the first but it is often difficult to see very fine tips. During this process some electrodes may become displaced, and eventually the relative positions must be determined. A device shown in *Figure 35.17* can be used.

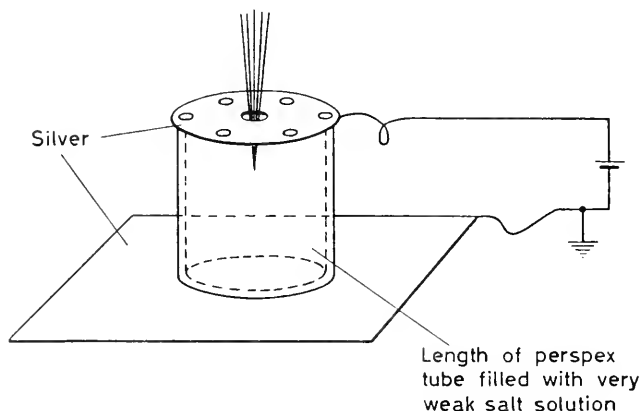


Figure 35.17 Fluid potentiometer for the determination of electrode spacing

A silver plate is placed at each end of a short tube, filled with a very dilute salt solution and 2 V applied between them. The electrode is introduced through a hole in the upper plate and the potential gradient is measured over a short distance which is determined with an accurate micrometer. The potential difference between neighbouring tips is used to find their spacing.

At the upper end of the electrode, electrical leakage presents difficulty. Electrode shanks of increasing length can be used to reduce this, and the

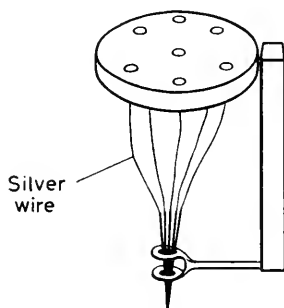


Figure 35.18 A method of holding a group of electrodes

orifices placed some distance apart. In addition the electrodes may be separated a little by interposing pieces of impregnated paper, and the glass coated with wax.

Such an assembly is cemented to a holder with wax (*Figure 35.18*), and thin silver wire introduced to establish connection with each. The electrodes can be filled with electrolyte solution in the usual manner.

REFERENCES

Transverse spacing of electrodes—Lateral spacing of electrodes is somewhat more difficult to achieve than longitudinal. The microcapillaries have to be fixed relative to each other with 'spacers' between or with wax. The electrodes are brought into place with the aid of a simple manipulator under a microscope. A manipulator which has been used is simply made of two L shaped pieces of Perspex, each provided with a vertical slide in addition, and grease between the sliding parts (Figure 35.19). Minute

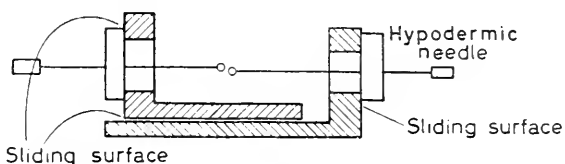


Figure 35.19 Section through a simple manipulator used for constructing a group of electrodes

quantities of Picene or other wax can be applied with the aid of a simple tool resembling a soldering iron. This consists of a thin wire soldered to a piece of thick wire which is surrounded by a small low voltage heating coil. The wax is applied while the tips are supported in place, as otherwise they are drawn together by the surface tension of the wax when fluid.

A combination of longitudinal and lateral spacing has been used to form different patterns of recording electrodes.

REFERENCES

- ¹ HODGKIN, A. L. and HUXLEY A. F. Resting and action potentials in single nerve fibres *J. Physiol.* 104 (1945) 176
- ² CURTIS, H. J. and COLE, K. S. Membrane resting and action potentials from the squid giant axons *J. Cell. Comp. Physiol.* 19 (1942) 135
- ³ LING, G. and GERARD, R. W. The normal membrane potential of frog sartorius fibres *J. Cell. Comp. Physiol.* 34 (1949) 382
- ⁴ BRITTON, H. T. S. *Hydrogen Ions* London; Chapman & Hall. 1955
- ⁵ CALDWELL, P. C. Measurement of intracellular pH by means of a tungsten microelectrode *J. Physiol.* 124 (1954) p 1P
- ⁶ ECCLES, J. C. *The Physiology of Nerve Cells* Oxford University Press. 1957
- ⁷ NASTUK, W. L. and HODGKIN, A. L. The electrical activity of single muscle fibres *J. Cell. Comp. Physiol.* 35 (1950) 39
- ⁸ MOREY, G. W. *The Properties of Glass* New York; Reinhold. 1938
- ⁹ HODGKIN, A. L. Personal communication (1957)
- ¹⁰ TEMPLETON, C. C. Drawing of small uniform glass capillaries by electrical beating *Rev. sci. Instrum.* 24 (1955) 880
- ¹¹ WEALE, R. A. A method for drawing microcapillary tubes *J. Physiol.* 112 (1950) 4P
- ¹² BROCK L. G., COOMBS, J. S. and ECCLES, J. C. The recording of potentials from motor neurones with an intracellular electrode *J. Physiol.* 117 (1952) 431
- ¹³ KAO, C.-Y. A method of making prefilled microelectrodes *Science* 119 (1954) 846
- ¹⁴ CRAIN, S. M. Resting and action potentials of cultured chick embryo spinal ganglion cells *J. Comp. Neurol.* 104 (1956) 285

GLASS MICROCAPILLARY ELECTRODES

- ¹⁵ ALEXANDER, J. T. and NASTUK, W. L. An instrument for the production of microelectrodes used in electrophysiological studies *Rev. sci. Instrum.* 24 (1953) 528
- ¹⁶ TASAKI, I., POLLEY, E. H. and OREGO, F. Action potentials from individual elements in cat geniculate and striate cortex *J. Neurophysiol.* 17 (1954) 454
- ¹⁷ CALDWELL, P. C. and DOWNING, A. C. The preparation of capillary microelectrodes *J. Physiol.* 128 (1955) 31 P
- ¹⁸ ECCLES, R. M. Intracellular potentials recorded from a mammalian sympathetic ganglion *J. Physiol.* 130 (1955) 572
- ¹⁹ FRANK, K. and FUORTES, M. G. F. Potentials recorded from the spinal cord with microelectrodes *J. Physiol.* 130 (1955) 625
- ²⁰ FATT, P. and KATZ, B. An analysis of the end plate potential recorded with an intracellular electrode *J. Physiol.* 115 (1951) 320
- ²¹ MARMONT, G. Studies on the axon membrane *J. Cell. Comp. Physiol.* 34 (1949) 351
- ²² FATT, P. and KATZ, B. The electrical properties of crustacean muscle fibres *J. Physiol.* 120 (1953) 171
- ²³ JANZ, J. G. and TANIGUCHI, H. Silver-silver halide electrodes *Chem. Revs.* 53 (1953) 397
- ²⁴ CASTILLO, del J. and KATZ, B. Local activity at a depolarized nerve muscle junction *J. Physiol.* 128 (1955) 396
- ²⁵ NASTUK, W. L. The electrical activity of the muscle cell membrane at the neuromuscular junction *J. Cell. Comp. Physiol.* 42 (1953) 249
- ²⁶ ADRIAN, R. H. The effect of internal and external potassium concentration on the membrane potential of frog muscle *J. Physiol.* 133 (1956) 631
- ²⁷ TASAKI, I. Properties of myelinated fibres in frog sciatic nerve and in spinal cord as examined with microelectrodes *Jap. J. Physiol.* 3 (1953) 73
- ²⁸ WOODBURY, J. W. Recording central nervous system activity with intracellular microelectrodes *Fed. Proc.* 12 (1953) 159
- ²⁹ WOODBURY, J. W. *J. Cell. Comp. Physiol.* 39 (1952) 323
- ³⁰ SOLMS, S. I., NASTUK, W. L. and ALEXANDER, J. T. Development of a high fidelity pre-amplifier for use on the recording of bioelectric potentials with intracellular electrodes *Rev. sci. Instrum.* 24 (1953) 960
- ³¹ McNICHOL, E. F. and WAGNER, H. G. A high impedance input circuit suitable for electrophysiological recording from micropipette electrodes *Nav. Med. Res. Inst.* 12 (1954) 97
- ³² HAAPANEN, L. and OTTOSON, D. A frequency compensated input unit for recording with microelectrodes *Acta Physiol. scand.* 32 (1954) 271
- ³³ BURCH, M. G. *Proc. Roy. Soc.* 48 (1890) 89
- ³⁴ LUCAS, K. On a mechanical method of correcting photographic records obtained from the capillary electrometer *J. Physiol.* 44 (1912) 225
- ³⁵ RUSHTON, W. A. H. *Proc. Roy. Soc. B* 123 (1937) 382
- ³⁶ LORENTE de No, R. A study of nerve physiology *Stud. Rockefeller Inst. med. Res.* (1947) 132
- ³⁷ TOBIAS, J. M. and BRYANT, S. H. An isolated giant axon preparation from the lobster nerve cord *J. Cell. Comp. Physiol.* 46 (1955) 163
- ³⁸ HODGKIN, A. L. and KATZ, B. The effect of sodium ions on the electrical activity of the giant axon of the squid *J. Physiol.* 108 (1949) 37
- ³⁹ TOMITA T. and FUNAISHI, A. Studies on intraretinal action potential with low resistance microelectrode *J. Neurophysiol.* 15 (1952) 75
- ⁴⁰ SVAETICHIN, G. Low resistance microelectrodes *Acta Physiol. scand.* 24 (1951) Suppl. 86.5
- ⁴¹ ROTHCHILD, W. L. The polarisation of a calomel electrode *Proc. Roy. Soc. B* 125 (1938) 283

REFERENCES

- ⁴² PHILPOT, St. J. L. Some experimental support for Stern's theory of the electrolytic double layer *Phil. Mag.* 13 (1932) 775
- ⁴³ FRICKE, H. The theory of electrolytic polarization *Phil. Mag.* 14 (1932) 311
- ⁴⁴ GRAHAME, D. C. The electrical double layer and the theory of electrocapillarity *Chem. Rev.* 41 (1947) 441
- ⁴⁵ WEALE, R. A. A new microelectrode for electrophysiological work *Nature, Lond.* 167 (1951) 529
- ⁴⁶ TAYLOR, G. F. A method of drawing metallic filaments and a discussion of their properties and uses *Phys. Rev.* (2nd series) 23 (1924) 655
- ⁴⁷ GRAY, J. A. B. and SVAETICHIN, G. Electrical properties of platinum tipped microelectrodes in Ringer's solution *Acta Physiol. scand.* 24 (1951) 278
- ⁴⁸ DOWBEN, R. M. and ROSE, S. E. A metal filled microelectrode *Science* 118 (1953) 23
- ⁴⁹ RENSHAW, B., FORBES, A. and MORRISON, B. R. Activity of isocortex and hippocampus electrical studies with microelectrodes *J. Neurophysiol.* 3 (1940) 74
- ⁵⁰ KENNARD, D. W. Local application of substances to the spinal cord *The Spinal Cord. Ciba Symposium.* London; Churchill. 1953
- ⁵¹ COOMBS, J. S., ECCLES, J. C. and FATT, P. The electrical properties of the motorneuron membrane *J. Physiol.* 130 (1955) 291
- ⁵² DEL CASTILLO, J. and KATZ, B. *Proc. Roy. Soc. B* 146 (1957) 339
- ⁵³ ELSON, L. Multibore glass sleeves for insulation *J. Sci. Instrum.* 30 (1953) 40
- ⁵⁴ VIS, V. A. A technique for making multiple bore microelectrodes *Science* 120 (1954) 152
- ⁵⁵ TOMITA, T. and TORIHANA, Y. Further studies of the intraretinal action potentials and on the site of the E.R.G. generation *Jap. J. Physiol.* 6 (1956) 119

OTHER ELECTRODES

I. A. SILVER

Electrodes that are used to stimulate or record electrical activity in biological fluids may become polarized unless special precautions are taken. The larger and more prolonged the current that passes between electrodes the greater the likelihood of polarization. When only small currents of short duration are being considered simple metal electrodes are adequate, but special modification of the electrode or stimulating apparatus or both may be necessary with higher current values.

Polarization appears as a partial blockage of current flow in conjunction with the presence of a back electromotive force. This back e.m.f. is associated with electrolysis of the medium around the electrode but the exact nature of over-voltage is not yet understood (see Butler¹; Kortüm and Bockris²). Although a study of polarization of electrodes in biological systems may be used in the investigation of the chemical state of the systems, most physiologists are chiefly concerned in avoiding or overcoming polarization phenomena.

NON-POLARIZABLE ELECTRODES

In experiments that require the passage of constant current or prolonged stimulation with monophasic shocks simple metal electrodes will polarize and the current passing will become progressively smaller, thereby altering the conditions of the experiment.

The general principle of 'depolarizers' is to avoid the electrolysis of the medium and in particular to suppress the production of bubbles of gas near the surface of the electrodes. This is most simply accomplished by keeping an electrode in contact with a saturated solution of one of its own salts.

The most generally useful non-polarizable electrode is the silver/silver chloride system, provided that only small currents are being employed. The electrode consists of a silver wire or plate on which silver chloride has been deposited electrolytically. The great advantage of the Ag/AgCl electrode is that it can be placed directly in contact with living cells, it is sufficiently strong to be easy to handle and insert deeply into tissues, and it is relatively non-toxic owing to the insolubility of silver chloride. A high concentration of silver ions is maintained around the electrode but there is very little diffusion and most tissues will tolerate the presence of silver electrodes for several weeks. Spermatozoa, however, are particularly susceptible to poisoning by silver ions so that investigations on electrical properties of semen should be conducted with electrodes of platinum or gold. Silver may also react with sulphides if it is left embedded in tissues for long periods.

PREPARATION OF NON-POLARIZABLE ELECTRODES

Platinum plated with platinum chloride is also commonly used and as platinum wire is stiffer than silver wire it is better for inserting deeply into resistant tissue. Platinum is however most frequently used without being plated with chloride since it has a very low gas over-voltage and is therefore virtually a non-polarizable electrode where very small currents are being considered.

The passage of heavier currents requires a metal electrode in contact with a very soluble salt if polarization is to be avoided. The zinc/zinc sulphate system is very commonly used in biological work, but as zinc is highly toxic it cannot be allowed to come into contact with the tissues directly. The greatest care is therefore necessary to confine the zinc ions behind a porous barrier, electrical contact being made through an agar/chloride bridge, and a cotton wick.

Mercury/mercuric chloride electrodes are also useful but, like zinc, mercuric ions are highly toxic to living tissues.

PREPARATION OF NON-POLARIZABLE ELECTRODES

Silver/silver chloride

For accurate and critical work it is necessary to plate silver electrodes very carefully. The following method produces stable electrodes with a large surface area.

Clean the silver wire with emery paper and degrease with Trilene or petroether. With the cleaned wire as the anode in N/10 HCl electrolyte and another silver wire as the cathode in circuit with a 2 V accumulator and a reversing switch, pass current through the circuit for 30 sec and then reverse the switch for 30 sec: repeat this three times. The anode-cathode distance should be very short and the plating dish should be in a dark box. After plating, the electrodes should be kept in Ringer's solution in the dark.

For very critical work the electrode can be imbedded in agar and connected to the tissues by a saline soaked wick.

Platinum/platinum chloride

Wash the platinum wire in distilled water followed by concentrated sulphuric acid. Rinse thoroughly in distilled water and place in platinum chloride solution. Connect the chosen electrode to the negative terminal of a 2 V accumulator and make a circuit through a dummy platinum electrode and a reversing switch. Pass current for 15 sec and then reverse it: repeat this 4 to 6 times ending with the chosen electrode positive, which will give a final even coat of platinum chloride. Wash the electrode and keep it in distilled water.

Zinc/zinc sulphate electrode

Ideally this consists of a pure zinc rod immersed in a saturated solution of zinc sulphate, but in practice zinc usually contains some impurities which are liable to produce small galvanic cells on the surface of the metal which will give rise to back e.m.f.s. and therefore such an electrode will appear to be polarized. This can be overcome by coating the surface of the zinc with a thin layer of mercury. The resulting amalgam behaves as a pure zinc surface.

OTHER ELECTRODES

An amalgamated zinc rod is placed in one arm of a *U* tube full of saturated zinc sulphate, the other arm being blocked by a ball of china clay soaked in saline. From the clay ball a cotton saline-soaked wick or a saline bridge with a wick makes contact with the tissues. The zinc ions are very mobile and will diffuse into the wick after about 24 hours so that new electrodes should be made up daily, otherwise tissue damage will occur.

SIMPLE METAL ELECTRODES

Although theoretically non-polarizable electrodes should always be used, in practice a great deal of work is carried out with simple wire electrodes. In the case of platinum, owing to its very low gas over-voltage, the plain metal is not polarized by very small short currents. Silver wire, as soon as it comes into contact with chloride-containing biological fluids, rapidly acquires a chloride layer and therefore is not easily polarized. Stainless steel and tungsten are chiefly used as micro- and semi-microelectrodes for recording action potentials from deep structures which can only be reached by electrodes with good mechanical properties. These two metals can be reduced electrolytically to very fine points comparable in size to glass pipette microelectrodes. Pure gold is difficult to handle owing to its softness but it may be used where an extremely inert metal electrode is necessary to check the performance of electrodes made of more reactive metals. It is particularly valuable in the measurement of oxidation-reduction potentials and oxygen tension in tissues.

The simplest forms of stimulating and recording electrodes consist of a suitable metal wire of appropriate thickness, insulated to within a short distance of the tip and fixed to some form of handle. Stimulating electrodes may be made in the form of a straight wire, a wire circle round the structure to be stimulated or, where large areas of stimulation are required, as a plate. A convenient form of electrode for stimulating or recording from whole nerve preparations can be made from a piece of Tufnol tubing fitted at each end with a plug of Perspex or sealing wax through which run the electrode wires.

It is frequently convenient to use unipolar electrodes as these can be made very small and there is only one possible stimulation point on the relevant structure. This system of course requires that there shall be somewhere in the preparation an indifferent electrode which is usually of a relatively large size giving a low current density. For acute experiments on whole animals a metal cylinder inserted into the animal's rectum makes a very efficient indifferent electrode, or a metal plate, usually of silver, can be placed subcutaneously. This latter method is also useful for chronic work. For measurement of oxidation-reduction potentials, pH and oxygen tension reference electrodes of simple silver wires plated with silver chloride and placed subcutaneously are quite satisfactory.

IMPLANTED ELECTRODES

Stimulation of, or recording from, deeply placed structures, in particular brain tissue, requires some modification of the simple metal electrode. The requirements for implanted electrodes are as follows: (a) they must be

IMPLANTED ELECTRODES

mechanically strong enough, and of suitable shape, to be pushed through overlying tissues; (b) they should be extremely well insulated except at the tip and designed so that they can be held in the clamp of some form of mechanical stereotaxic instrument in order that they can be placed accurately in position in the depths of the tissue.

Platinum or silver wire insulated with glass capillary tubing and shellac or synthetic resin are commonly used for stimulating electrodes (Harris³), the glass providing both the insulation and the rigidity. This type of electrode will penetrate brain tissue readily but is easily damaged by contact with bone.

To avoid damage to the electrode which may occur if it encounters an unexpected resistance the wire should be mounted on a sprung platform which can be watched during insertion (Figure 36.1). Unusual stress on

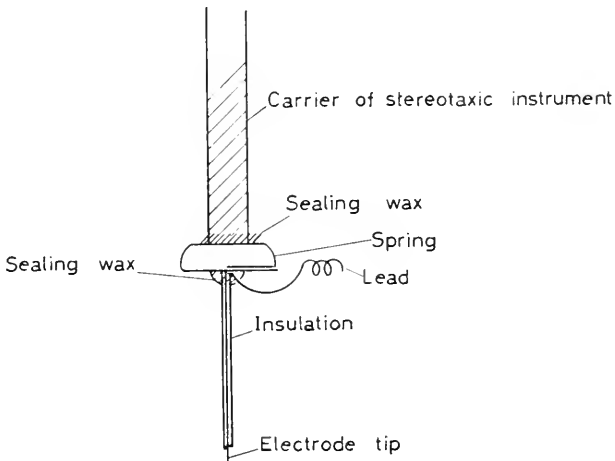


Figure 36.1 Sprung electrode base

the electrode will bend the spring before the electrode breaks and this is particularly useful if the tip of the electrode is situated near the skull floor.

When electrodes have to be implanted and left *in situ* for long periods, for example during investigations of the effect of stimuli on conscious animals, the reaction of the tissues to the presence of the electrode must be considered. Any strong tissue reaction will cause the electrode to become surrounded by inflammatory cells which will alter markedly the relationship of the electrode to the normal tissues. Chronically implanted electrodes should be as far as possible both chemically and mechanically inert. Platinum and gold satisfy these requirements but gold is too soft for general purposes. Most chronically imbedded electrodes are made of platinum or platinum-iridium alloy, although stainless steel, tungsten and silver are also employed. Silver may give rise to black deposits in the tissues after several weeks (Harris³) which are probably silver sulphide.

The surface of an implanted metal electrode should be carefully polished before insertion to eliminate roughness, and the insulating material must be non-toxic and as far as possible non-irritant to tissues. It is an added

OTHER ELECTRODES

advantage if the insulator is 'non-wettable', but if it is not, the insulator is best coated with silicone.

FIXATION OF ELECTRODES

Where electrodes are to be left in place over days or weeks it is vital that the exact position of the tip be known, and also that it should stay firmly fixed in position. A poorly fixed electrode will not only give useless information but may also destroy large areas of brain.

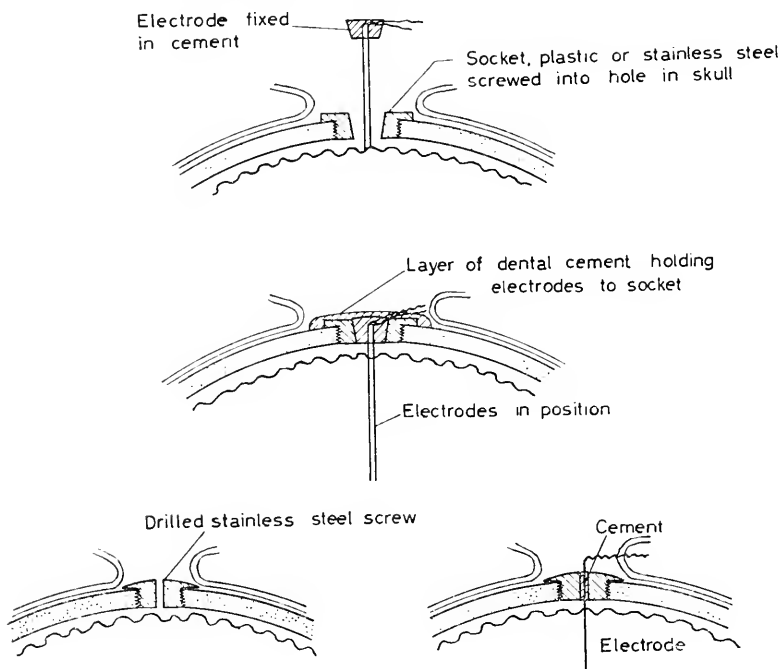


Figure 36.2 Fixation of electrodes in brain

A simple technique for implanting electrodes in the brain of rabbits is described by Harris³ and in cats by Bradley and Elkes⁴. The electrode, either unipolar or bipolar, is made of suitable wire insulated with non-irritant resin or glass. It is held in a stereotaxic instrument and inserted into the brain through a hole drilled in the roof of the skull and an incision through the dura mater. By using known co-ordinates for the particular brain the electrode tip can be placed with considerable accuracy in any desired position, which is then checked by radiograph before the electrode is finally fixed. The electrode and hole in the skull are packed with dental acrylic resin which hardens in a few minutes leaving the electrode firmly held in the skull. The wire is then released from the clamp, any excess is cut off and the top with its connecting lead is smoothed over with more resin. The skin can be sutured over the top of the electrode base and the leads taken subcutaneously to any desired point before they are brought

FIXATION OF ELECTRODES

to the surface. The details of the technique are shown as *Figures 36.2 and 36.3*.

With unipolar electrodes a suitable indifferent electrode can be made from a silver plate placed subcutaneously over the anterior part of the

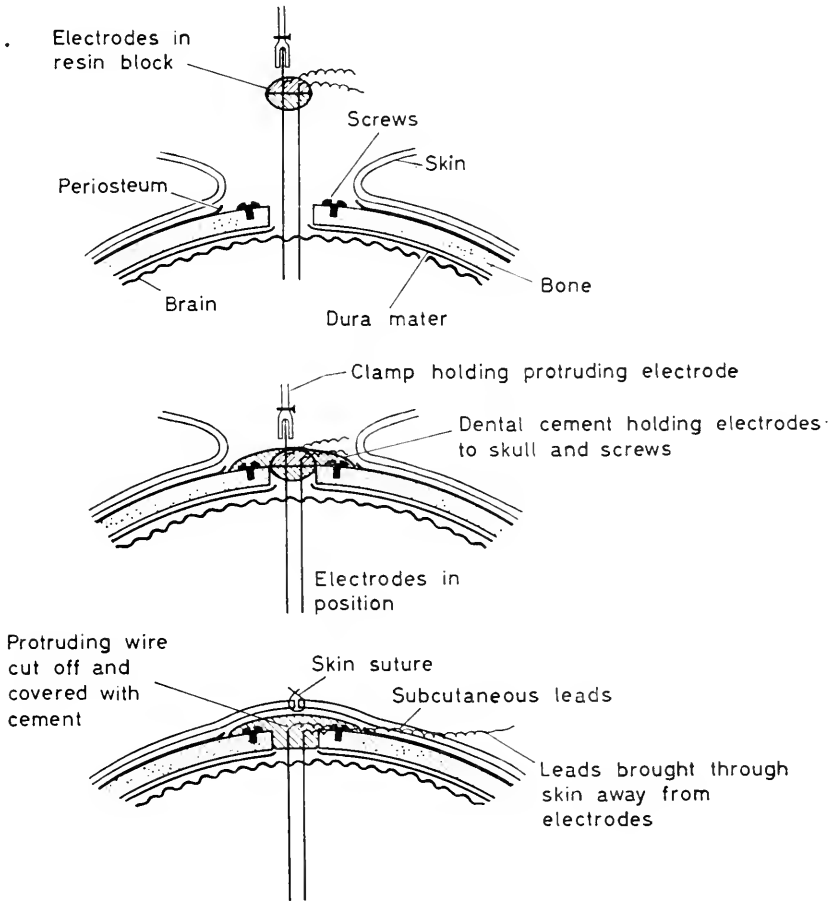


Figure 36.3 Fixation of electrodes in brain

skull. It is important that the plate should not be placed on the posterior of the head as the movements of muscles in that region can cause displacement of the plate which may damage both the leads to the electrode and the animal.

Suitable insulating material for imbedded electrodes includes most of the modern synthetic resins with the required electrical properties except the phenol formaldehyde group, many of which give rise to tissue reactions if left in the body for long periods. The polyester resins should be mixed with a minimum of accelerator and catalyst if they are to be used for insulating or 'potting' small electrical components that are to be buried

OTHER ELECTRODES

in animal tissue, since excess of the commonly used catalysts and accelerators may be toxic. Polyesters have the advantage over most other resins in that they can be both transparent and cold setting, and either hard and rigid or soft and pliable when polymerized.

A difficulty common to all electrodes is that of determining where the tip has been during stimulation or recording. The smaller the electrode the more difficult the task of tracing the track through the tissues during histological reconstruction after an experiment. If the electrode has been used in one place only the tissue around it should be fixed by perfusion while the electrode is *in situ* as this will leave the electrode track clearly visible under the microscope. This has the added advantage that an estimate of the degree of shrinkage of the tissue during fixation can be made by measuring the length of the electrode track before and after fixation. If many tracks have been made and if it is necessary to follow each one some sort of marker system is essential. In the case of relatively large electrodes there will usually be a small amount of haemorrhage which will render the tracks visible in sections, but in many cases, and particularly where very small electrodes are used, no such convenient marking takes place. Before insertion an electrode may be coated with some form of pigment such as carbon, which will be left adhering to the track walls, but results with this method are very variable and may be misleading as to the position of the tip of the electrode. Small electrolytic lesions can be made around the tip of some electrodes at the end of an experiment by passing a small direct current through the electrode which will give a general indication of the position of the point of the electrode in relation to surrounding nerve tracks and nuclei. If steel wires are being used iron can be deposited electrolytically in the tissue from the point of the electrode and this is detected later by a histochemical Prussian blue reaction (Marshall⁵, Adrian and Moruzzi⁶).

After using microelectrodes it is frequently impossible to identify the track or position of the tip in a histological reconstruction of the tissue. It is possible that if such an electrode were coated before insertion with some form of radioactive dust or colloid, sufficient radioactive material might be transferred to the tissues to be detectable afterwards by autoradiographs. ¹⁴C might well be a suitable 'marker' substance.

METAL MICROELECTRODES

Metal electrodes with very fine tips for recording from very small areas or even from single cells have the advantage over glass pipette microelectrodes of being relatively robust, and can be used for both acute and chronic work. In all other respects however they are generally inferior to glass electrodes for critical work.

Steel microelectrodes can be made from ordinary darning needles but are better produced from a good stainless steel wire such as Staybrite. They should be used in conjunction with some form of mechanical manipulator as the fine tip is very fragile, but it is possible to leave them imbedded in brain tissue for long periods provided they are firmly fixed in position. They can be made relatively easily with tips down to 10 μ diameter and

with greater care as small as $1\ \mu$. These very small tips are not very satisfactory as they are extremely brittle and usually of very high impedance.

The method of preparation is simple but care is needed to obtain really fine regular points (Bishop and Collier⁷, Grundfest, Sengstaken, Oettinger and Gurry⁸). A steel wire is connected in circuit with a copper plate (negative) in a Petri dish of N hydrochloric acid to a 6 V accumulator. The needle tip is dipped into the dish for 2 minutes and then withdrawn, dipped again, withdrawn and so on. As the process continues the battery voltage should be decreased to 2 V. When a fine enough point has been obtained the needle is disconnected, washed in water and degreased in acetone. When dry it is insulated with Synobel varnish (I.C.I. Ltd.) by dipping the whole electrode into a deep pot of the varnish and withdrawing it slowly. The needle is then turned with its point upwards and allowed to drain, after which it is baked in an oven at 200°C . Several coats of varnish are necessary but even so the extreme tip of the needle often remains uninsulated so that the electrode is at once ready for use. If however the tip is insulated it is only necessary to touch it with a piece of paper to break off the insulation.

37 s.w.g. Staybrite steel wire (Johnson, Matthey and Co., London) may be used if the electrodes are to remain in the tissues for more than a few hours, and this necessitates a minimum of 4 V for electrolytic reduction of the tip.

A tungsten microelectrode has recently been described by Hubel⁹ which is said to behave as well as microglass electrodes. The advantage of tungsten over steel is that it is much stiffer and therefore the original wire can be very thin but still have good mechanical properties. A tip less than $5\ \mu$ in diameter is suitable for picking up extracellular potentials but for intracellular work the tip must be smaller than $1\ \mu$.

Tungsten wire $125\ \mu$ diameter is inserted into a 27 gauge hypodermic needle which is then crimped to hold the wire tightly. The short end of tungsten protruding from the point of the needle is then reduced by electropolishing to the required diameter. The electrolysis is carried out in a cell containing saturated aqueous potassium nitrite using 2–6 V, a.c., and a nearby carbon electrode. The wire should be raised and lowered as described in the preparation of steel microelectrodes so that a taper is produced on the wire. If raising and lowering is not carried out a very abrupt point will be produced which will not penetrate tissues very readily. When a suitable taper has been achieved the last part of the reducing process should be completed without moving the electrode in the cell in order to make a very fine uniform point. Tips as small as 0.5 to $0.05\ \mu$ diameter are said to be obtainable without great practice or skill. The wire should be washed in detergent and after drying coated with clear vinyl lacquer which has been allowed to thicken slightly by exposure to air. The coating is carried out by dipping the electrode into a deep pot of lacquer then withdrawing it and allowing it to drain tip downwards, after which the tip only is continually re-dipped until no further beads of lacquer form on it. It is then dried for 24 hours.

The insulation can be tested for leaks by placing the electrode in 0.9 per cent NaCl and making it the cathode in circuit with a 12 V accumulator.

OTHER ELECTRODES

Any break in the insulation will be detectable by the bubbles of gas which will rise from it. This method of testing is useful for any resin-insulated metal electrodes. For testing insulation on very fine intracellular electrodes it is convenient to measure the resistance of the electrode as it is lowered into a saline bath. Any abrupt change in resistance is an indication of a flaw in the insulation.

Tungsten microelectrodes of this type have a resistance of 25 to 200 M Ω in different electrodes at low frequency (100 c/s) and low current density. At higher frequencies (5–10 kc/s) the impedance falls to values around 0.5 M Ω .

CONCENTRIC ELECTRODES

These are useful for recording from or stimulating relatively small areas of tissue in whole animals which may be conscious or anaesthetized. Their chief application is in the field of electromyography where they can be used to detect activity in motor units for clinical or experimental purposes. They are very robust and can be inserted into dense tissues through intact skin. A simple concentric electrode can be constructed as follows:

Pass about 70 cm of Eureka resistance wire (gauge 44) through a 26 gauge Solila type hypodermic needle of the desired length so that about 10 cm of the wire projects from the point of the needle and about 60 cm from the back of it. Dip the protruding 10 cm of wire into Damarda lacquer (Grade L 3128, Bakelite Ltd.) thinned with equal proportions of Bakelite S 11210 thinner. Pull the needle over the lacquered wire until only about 0.5 cm protrudes from the tip and then place in an oven at 150°C for 3–4 minutes. Too hot an oven will cause bubbling of the lacquer and roughness on the electrode. Solder another piece of wire on to the base of the needle, degrease in Trilene and dip needle into lacquer. Leave, point upwards, to drain for 5 minutes which will allow any bubbles to flow to the base of the needle, and bake for 3 to 4 minutes. Repeat the degreasing and lacquering five or six times. The tips of the wire and needle are then ground flush on a fine hard stone and the electrode tested for insulation and conduction.

ELECTRODES FOR MEASURING OXYGEN TENSION

If a small potential of the order of 0.6 V is applied to a noble metal electrode in animal tissue the current that flows is approximately proportional to the oxygen concentration in the tissue (Davis and Brink¹⁰, Cater, Phillips and Silver¹¹). Two types of electrodes can be used for such measurements: (a) the simple 'open ended' electrode which is merely a gold or platinum wire insulated with resin and ground to a flat end, or (b) the 'recessed' electrode which consists of platinum wire fused into glass capillary tubing so that the tubing projects beyond the metal (*Figure 36.4*).

The open-ended electrodes are useful for making continuous records of changes in oxygen tension while the recessed type will make intermittent readings of absolute oxygen concentration.

Currents of the order of 10⁻⁸ to 10⁻¹⁰ amps are obtained depending on the size of the electrode and the oxygen tension in the tissue.

ELECTRODES FOR MEASURING OXYGEN TENSION

Open-ended electrodes can be made with platinum, gold or iridium-platinum wire of any desired size down to $25\ \mu$ (iridium/platinum) but gold appears to give the most consistent results although it is too soft for easy insertion into tough organs. Insulation with the epoxy resin Araldite

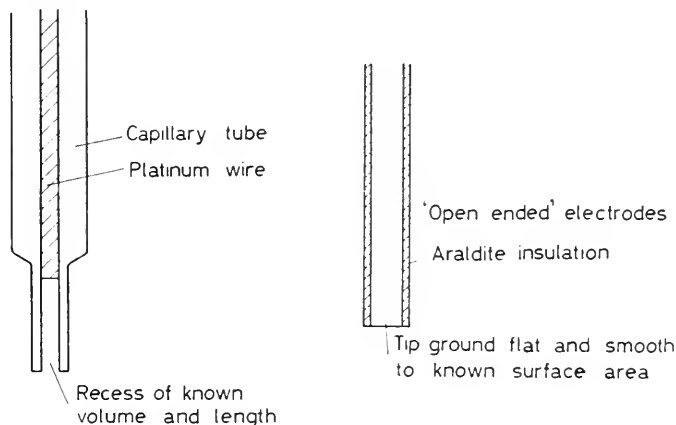


Figure 36.4 Oxygen tension electrodes

985 E (Aero Research Ltd., Duxford) is the most satisfactory as this resin has very good insulating properties, is extremely tough and is slightly flexible. Several coats are necessary and each must be baked for about 10 minutes at 200°C so that it is quite hard before the next layer is applied. The insulated wire is placed in a Tufnol clamp and ground flat at the end by hand, using a very fine hone. Mechanical grinding, e.g. with a jeweller's wheel, is not satisfactory as it tends to spread a thin layer of metal over the

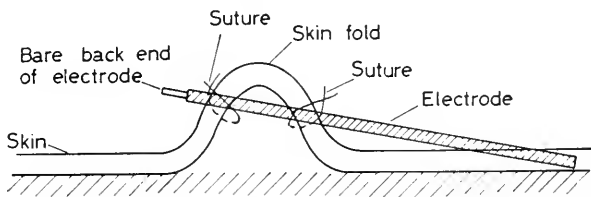


Figure 36.5 A method of fixing an intradermal or subcutaneous electrode

end of the insulation, thereby increasing the effective area of the exposed end of the wire. It is essential that the electrode tip should be ground quite flat and have a known area if approximately absolute levels of oxygen tension are being measured at the same time as changes. These electrodes are relatively robust and may be inserted into the tissues by hand or mechanically. If changes over several days are to be measured they may be left imbedded in the animal with only a small part projecting through the skin. This may be useful in following the development of inflammatory reactions (Figure 36.5).

Recessed electrodes are made with platinum wire which is first degassed by heating to white heat to prevent bubble formation in the next stage.

OTHER ELECTRODES

The degassed wire is fused into a soft glass capillary tube so that the tube projects beyond the end of the wire in a regular cylinder for a known distance; the glass then being annealed to prevent cracking. Alternatively the platinum can be fused into the tube so that it is flush with the end, or even projects, after which the recess is produced by dissolving the metal to the required depth with aqua regia. These electrodes are very delicate and can only be used for acute experiments. They are more suitable for measurements at surfaces than in the depths of tissues. At a given oxygen tension the current passing through a recessed electrode remains constant over a range of applied potentials of approximately 0.2 to 0.8 V, but the open-ended type do not behave in this convenient way so that the current varies with the voltage without any simple relationship between the two. Oxygen tension electrodes sometime suffer from so-called 'poisoning' after being exposed to the tissues for several hours. This appears as a progressive fall in current when the oxygen tension remains constant. It appears to be due to tissue fluid proteins or bacterial activity on the surface of the metal. If electrodes are carefully cleaned and sterilized by boiling before insertion, and are implanted with aseptic precautions 'poisoning' can usually be avoided. Reference electrodes for use in oxygen tension measurements can be silver/silver chloride for imbedding or calomel half-cell with a bridge for surface contact.

Electrodes similar to the open-ended type described above can be used for measurements of oxidation-reduction potentials in animal tissues, with silver/silver chloride as the reference electrode. For a balanced electrometer valve circuit suitable for multi-channel measurement of redox potentials see Cater, Phillips and Silver^{11,12}.

Many other types of electrodes are available for measuring oxygen tension, the most accurate of which is the dropping mercury electrode, but this is unsuitable for use on biological material except for suspensions of cells, and even here the toxicity of the mercury renders results obtained with living material rather open to criticism (Petering and Daniels¹³).

GLASS ELECTRODES FOR MEASURING pH IN TISSUES

Several types of electrodes have been used in an attempt to measure intra- and extracellular levels of pH in living preparations. A micro-tungsten electrode described by Caldwell¹⁴ was affected by factors other than pH, such as cystein, H₂S, calcium ions and oxidizing agents. Antimony electrodes are sensitive to changes of oxidation-reduction potential and oxygen tension. Glass electrodes however appear to be affected only by changes in pH in the range pH 0 to pH 9 (Dole¹⁵) so that these are the electrodes of choice for pH work despite the disadvantage of extreme fragility.

For work on animal tissues a glass electrode must be sufficiently small to be entirely imbedded in the tissue concerned, of suitable shape for insertion and must have extremely good insulation near its tip. It should give a range of at least 50 mV per pH unit and have a resistance in the range of 50 to 1,000 M Ω depending on the type of electrode. The smaller the tip the greater the impedance.

GLASS ELECTRODES FOR MEASURING pH IN TISSUES

Glass electrodes for physiological use are usually based on the capillary electrode designed by Voegtlin, Kahler and Fitch¹⁶, and modified for the particular purpose required. An extracellular electrode suitable for most purposes where pH changes in small blocks of tissue are to be measured is described by Sonnenschein, Walker and Stein¹⁷. A piece of Corning 015 glass capillary tubing is cleaned in nitric acid and washed in water. It is drawn down in a microflame to an outside diameter of about 2 mm. This narrow tip is re-drawn to 0.2 to 0.5 mm and the end is sealed very carefully so that no globule of glass forms, and no pinpricks are left in the seal.

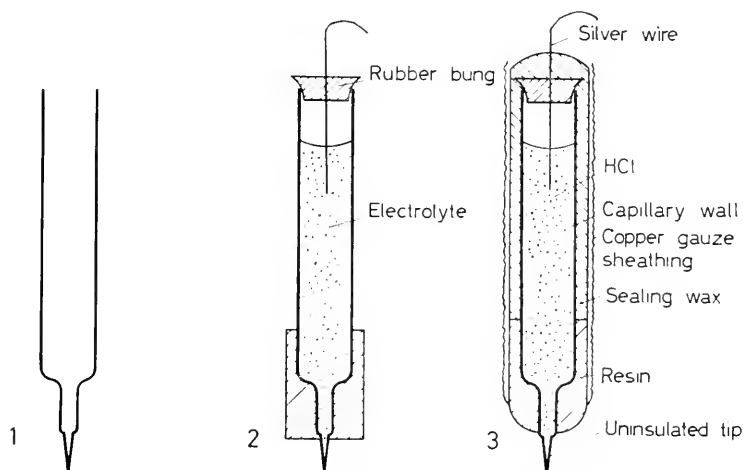


Figure 36.6 Extracellular pH electrode

The tube is then half filled with either a saturated solution of silver acetate in 50 per cent acetic acid or with 0.1 N hydrochloric acid. With the first solution a pure silver wire is needed as the inner electrode but with the second electrolytically plated silver/silver chloride is necessary. Air bubbles must be eliminated and then the electrode is left in boiling water for 2 hours as this equilibrates the asymmetry potentials of the glass more quickly. After boiling the tube is placed in warm air at 150°C so that the outside of the glass becomes heated. Apply a good insulating resin which will harden immediately on cooling. Sonnenschein *et al.*¹⁷ recommend de Khotensky cement for the critical region which consists of the 2 mm diameter part of the electrode and the first part of the main tube (Figure 36.6). Fill the tube with the rest of the chosen electrolyte and seal it with a rubber bung through which projects the internal electrode of silver wire. The main part of the glass is then covered with a layer of sealing wax which is continued over the rubber bung. A final coat of paraffin wax is added to everything except the very fine tip of the electrode, and the whole of the insulated part is then covered with fine copper gauze sheathing. The complete electrode should be kept in distilled water or buffer at pH 6.

The resistance of this type of electrode is about 400 MΩ and it should

OTHER ELECTRODES

give a range of 55 to 60 mV per pH unit. The temperature sensitivity is of the order of 2 to 3 mV per degree centigrade.

An intracellular electrode for large cells (Caldwell¹⁴) can be made on the same general pattern (Figure 36.7). The details are as follows: Low

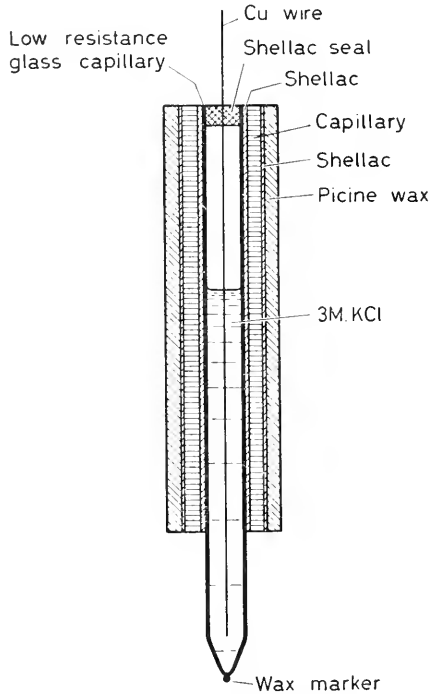


Figure 36.7 Intracellular pH electrode

resistance tubing (A. W. Dixon and Co., Beckenham, Kent) of 0.5 cm external diameter is pulled down and thinned to a capillary of 50–80 μ diameter. Coat the tube with shellac to within 1.5 cm of the tip and then place it inside another capillary tube of internal diameter 80–100 μ so that only the uninsulated part of the smaller capillary protrudes from the end of the larger one. Shellac both tubes together avoiding the tip of the internal capillary, and then cover the larger capillary with Picine wax. Fill the tip of the inner tube with 0.1 N HCl and seal off in a microflame until only 0.5 cm projects from the insulated part. A 30 μ copper wire is then inserted into the open back end of the inner tube so that it reaches the acid and is then sealed into position with shellac. The very tip of the capillary can be coated with a minute spot of Picine wax to make it more easily visible if desired.

These micro-glass electrodes have a very high resistance of the order of 1,000 M Ω and give a range of 50 to 60 mV per pH unit. If the electrode has a greater resistance than 1,000 M Ω or if it gives less than 50 mV per pH unit it should be discarded.

The insulation of pH electrodes near the tip is extremely critical. This

REFERENCES

can be tested for any type of capillary electrode by having a tube of agar at a known pH with a layer of fluid of a different pH above it. The electrode is lowered into the tube so that the uninsulated tip is just buried in the agar while the insulated part is in the fluid which is added when the tip is well imbedded. The reading of pH should correspond to that of the agar only if the insulation is complete. As the electrode is withdrawn the reading should change immediately to the value for the fluid.

REFERENCES

- ¹ BUTLER, J. A. V. *Electrocapillarity* London; Methuen. 1940
- ² KORTÜM, G. and BOCKRIS, J. O'M. *Textbook of Electrochemistry* Amsterdam; Elsevier. 1951
- ³ HARRIS, G. W. *Phil. Trans. Roy. Soc.* 232 B (1947) 385
- ⁴ BRADLEY, P. B. and ELKES, J. *EEG. Clin. Neurophysiol.* 5 (1953) 451
- ⁵ MARSHALL W. H. *Stain Tech.* 15 (1940) 133
- ⁶ ADRIAN, E. D. and MORUZZI, G. *J. Physiol.* 97 (1939) 153
- ⁷ BISHOP, P. O. and COLLIER, R. *J. Physiol.* 112 (1951) 8P
- ⁸ GRUNDFEST, H., SENGSTAKEN, R. W., ÖETTINGER, W. H. and GURRY, R. W. *Rev. Sci. Instrum.* 21 (1950) 360
- ⁹ HUBEL, D. H. *Science* 125 (1957) 549
- ¹⁰ DAVIES, P. W. and BRINK, F. *Rev. Sci. Instrum.* 528 (1942) 13
- ¹¹ CATER, D. B., PHILLIPS, A. F. and SILVER, I. A. *Proc. Roy. Soc. B* 146 (1957) 289
- ¹² CATER, D. B., PHILLIPS, A. F. and SILVER, I. A. *J. Physiol.* 129 (1955) 33P
- ¹³ PETERING, H. G. and DANIELS, F. J. *J. Amer. chem. Soc.* 60 (1938) 2796
- ¹⁴ CALDWELL, P. G. *J. Physiol.* 126 (1954) 169
- ¹⁵ DOLE, M. *The Glass Electrode* New York; Wiley. 1941
- ¹⁶ VOEGTLIN, G., KAHLER, H. and FITCH, R. H. *Nat. Inst. Hlth. Bull.* 164 (1935) 15
- ¹⁷ SONNENSCHIN, R. R., WALKER, R. M. and STEIN, S. N. *Rev. Sci. Instrum.* 24 (1953) 702

PART IV
COMPLETE APPARATUS

POWER PACKS

A power pack is that part of an electronic equipment which, drawing energy from the electric supply mains, is responsible for supplying appropriate power to the remainder. A complete apparatus may draw all its power from a master power pack—or 'power unit'—or the several units comprising the apparatus may each be endowed with their own, individual small power packs. The latter scheme is conducive to greater flexibility, but if high grade power supplies are necessary it is usually uneconomic.

Three types of supply are usually involved:

(1) *HT supplies for valves*—Invariably a positive supply, of a few hundred volts above earth, and very frequently—particularly with direct-coupled and pulse apparatus—a negative supply, a few hundred volts below earth.

(2) *LT supplies for valve heaters*—6.3 V a.c. or d.c. depending on the application, is the usual heater rating for general electronic work.

(3) *EHT supplies* of a few thousand volts, most particularly for cathode ray tubes, but also for Geiger-Müller tubes and Multiplier Photo Cells.

These three types of supply will now be discussed in detail.

HT SUPPLIES

Unstabilized supply

The most elementary type of HT supply is shown, in its commonest form, in *Figures 37.1* and *37.2*, using thermionic and non-thermionic rectification respectively. A full-wave capacitor-input rectifier circuit feeds a simple smoothing filter comprising a single choke and capacitor; the performance of such a filter has been discussed in Part I; typically the ripple on the output is of the order of a volt. In the thermionic case the valve manufacturer sometimes specifies the use of series anode resistors as shown, to limit the valve current during the switching-on surge. Notice the protective fuses, also the primary fuse and main switch connected to the phase, not the neutral, side of the mains. In laboratories still fitted with two-pin plug and socket distribution systems, where the sense of connection can be reversed, double-pole switching and fusing is advisable.

A simple unstabilized supply of the above kind has the disadvantages that its output depends on the mains voltage, and that its output impedance is rather high, a few hundred ohms; that is, the regulation is poor. It would be quite suitable for feeding, say, a small loudspeaker amplifier; the latter draws, on the average, a constant load current (the output valve being in class A) so that the high output impedance does not matter, and fluctuation in HT voltage caused by fluctuating mains voltage might produce small fluctuations in gain, but these would be quite unimportant. It would not be

POWER PACKS

so successful at supplying, say, a double-channel stimulator; the sort of distressing effect one would find would be that turning up the amplitude control of output number 1 produces the expected effect, but also slightly reduces the amplitude of output number 2. This results from a reduction

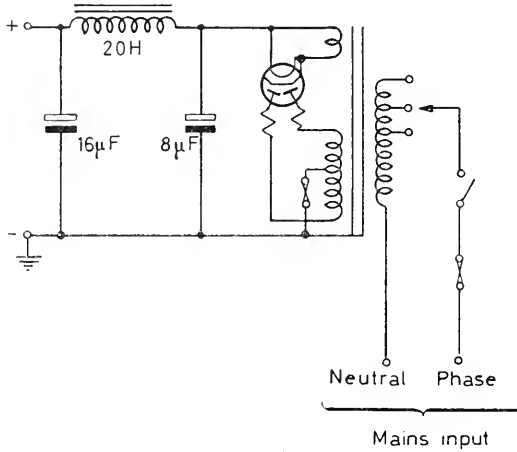


Figure 37.1

in HT voltage caused by the higher current drain of the No. 1 channel, and would be mitigated if the power pack regulation were better, i.e. if the HT were stabilized. It is possible to use stimulators fed from unstabilized supplies, but not very convenient; none of the controls can be properly

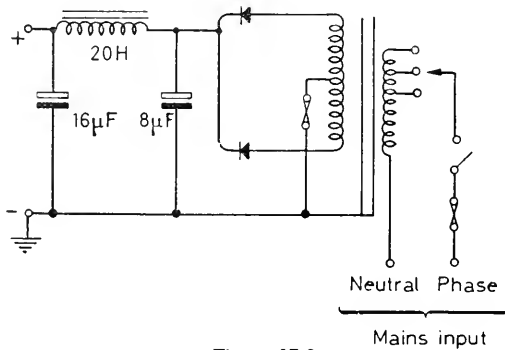


Figure 37.2

calibrated. If financial stringency dictates the use of simple power packs in instrument applications, then the output impedance should be kept as low as possible by using a low resistance choke, and rectifiers rated for a rather higher current than necessary. The extremely low forward resistance of germanium and silicon junction rectifiers would also be a help here, but as yet they are not very cheap.

Cold-cathode-diode-stabilized HT supplies

The design of cold-cathode-diode circuits has been discussed in Part I and a complete HT supply stabilized in this manner has the appearance of

HT SUPPLIES

Figure 37.3, which also provides a negative supply from the same transformer. The use of cold-cathode regulators will considerably reduce the effect of mains voltage fluctuations on the output, but is not conspicuously effective in reducing the output impedance, which will still be of the order of

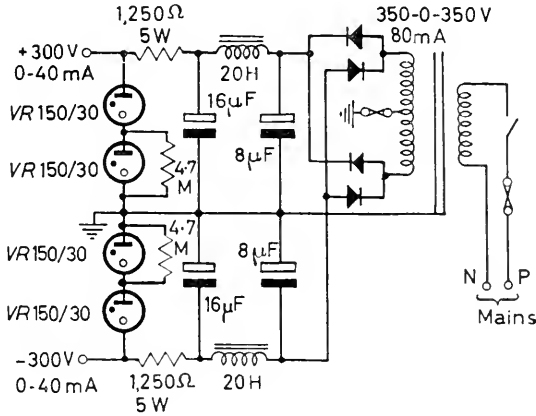


Figure 37.3

hundreds of ohms. It is not possible to be more precise than this because individual tubes vary appreciably in their ability to maintain a near-constant potential difference over a wide range of current. Cold-cathode-diode-stabilized power packs are therefore suitable for applications where the load is rather constant, such as precision oscillators, differential amplifiers, etc.

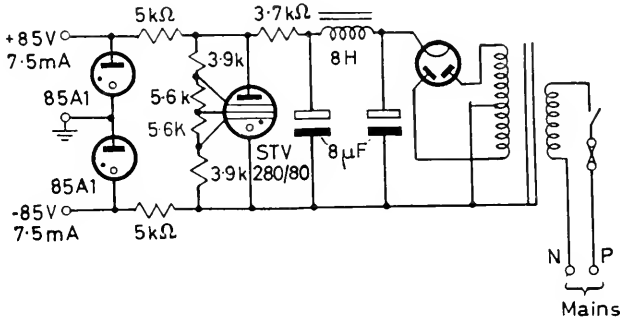


Figure 37.4

Even greater independence of mains voltage fluctuations may be had by connecting stabilizer circuits in cascade, as for example in the arrangement in Figure 37.4, devised by Luke¹ for supplying E.E.G. amplifiers.

Electronically stabilized HT supplies

In order to secure both independence of the mains voltage and a really low output impedance it is necessary to use electronic stabilization, employing hard valves. Innumerable designs for doing this have been produced,

but the principle of them all is the same. The output voltage, or a fixed fraction of it, is compared with a reference voltage derived from a battery or, more commonly, a voltage reference tube. Any difference between them is amplified and used to correct the output voltage by altering the d.c. resistance of a valve either in series with the output terminals (series stabilization) or in parallel with them (shunt stabilization or 'absorber valve'). Series stabilization is commoner and will be dealt with in greater detail.

In its simplest form the electronic series voltage stabilizer is shown in *Figure 37.5*. V_1 is a high μ triode, and V_2 is the series control valve. VR is

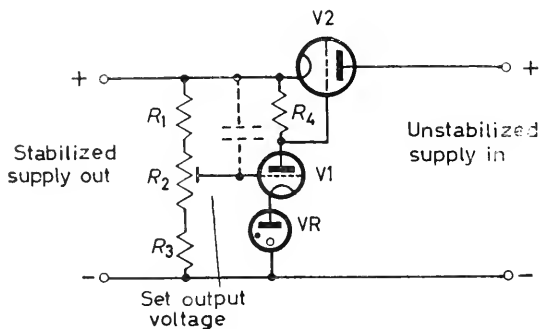


Figure 37.5

the voltage reference tube. R_1 , R_2 and R_3 form a potential divider across the output, from which a pre-set fraction of the output voltage is applied to the grid of V_1 . R_4 is the anode load for V_1 , and current flows down R_4 , through V_1 and VR , igniting VR and maintaining V_1 cathode at a fixed potential above earth. When the circuit is in equilibrium the setting of R_2 is such that V_1 grid is slightly negative to its cathode. If, for some reason, the output voltage tends to rise, V_1 grid goes positive, increasing V_1 anode current and causing V_1 anode (and therefore V_2 grid) to fall in potential. This increases the resistance of V_2 and so opposes the original tendency—and vice versa.

Clearly this device not only opposes slow changes in the output, but also opposes ripple. Electronically stabilized power packs are characterized by very low ripple in the output despite quite low value chokes and smoothing capacitors. Ripple is further reduced by connecting a capacitor (shown dotted) as in *Figure 37.5*, since by this means nearly all the ripple signal is applied to V_1 grid to produce corrective resistance changes in V_1 , instead of just the fraction determined by the potential divider $R_1R_2R_3$.

This simple electronic stabilizer works quite well, reducing the output impedance to a few ohms, but, operated as it is by its own error, it can never give perfect stabilization. An improvement due to Lindenhovius and Rinia² is to supplement the signals applied to V_1 grid with one derived from the unstabilized input, to help correct for mains fluctuations, and one derived from the output current, which reduces the output impedance. The first signal is applied via R_5 , and the second is developed across R_6 , to give the circuit of *Figure 37.6*. In this manner the performance can be improved but only over a rather restricted range of output current.

HT SUPPLIES

Alternative methods of improving stabilizer performance revolve round the necessity to get the maximum possible amplification from V_1 . One approach lies in the use of positive feedback; this is done in the circuit of *Figure 37.7*, which was shown to the author by J. A. Popple. The method seems to originate in a paper by Gray³. A definite fraction of the output

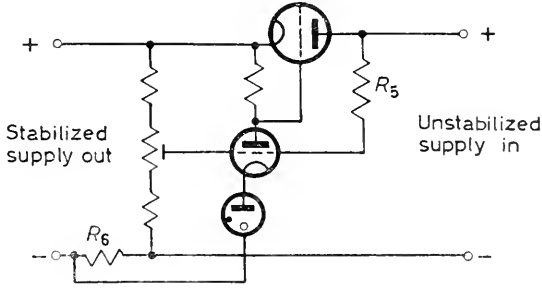


Figure 37.6

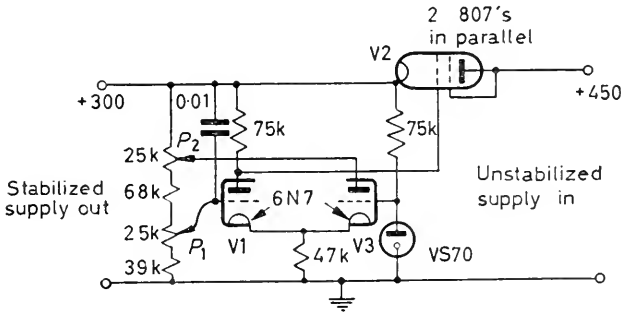


Figure 37.7

is applied via P_1 to V_1 , which controls V_2 in the same way as in *Figure 37.5*. The reference potential is applied not direct to V_1 cathode but to V_3 grid. V_1 and V_3 form a differential amplifier with the 47k resistor as common 'tail'. Clearly when P_2 slider is at the top there is no positive feedback, but as P_2 slider is moved down feedback comes into action; for an increase in output voltage raises V_1 grid and hence the common cathode potential. V_3 anode current therefore falls, reducing the voltage drop across that fraction of P_2 above the slider and hence increasing the slider potential. This increase in potential must be communicated down the chain to V_1 grid, hence the feedback is positive.

In use, P_1 is used to set the output voltage to the required value, whilst P_2 determines the output impedance. After altering P_2 , P_1 will require re-adjustment. The output impedance can be made zero, or even negative, but instability can occur if this is done.

An important advantage possessed by this circuit, and not by those of *Figures 37.5* and *37.6*, is that the use of differential amplification between the reference potential and P_1 slider confers the usual advantage that the circuit is relatively unaffected by the variations in cathode emission which follow from feeding V_1 from an unstabilized 6.3 V a.c. supply.

POWER PACKS

Another important improved form of voltage stabilizer is due to Attree⁴, who points out a basic difficulty in the design of stabilizers of the classical type; this is that since the optimum mean bias in the series control valve determines the permissible standing voltage drop across the load of the amplifier valve, and since this drop is not as large as one would like, one is obliged to make shift with either a low value of load resistance for the amplifier valve or a low anode current. With valves the mutual conductance falls at low anode currents. In either case the gain is poor. Attree goes on to show how a slightly modified cascode circuit overcomes this difficulty, enabling him to quote an output impedance for his circuit of only 0.2 ohms, and a ripple and noise content of less than 500 μV . The

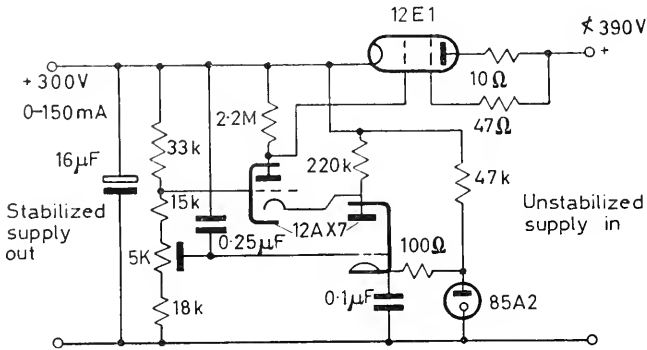


Figure 37.8

modification consists merely in by-passing the upper cascode triode by a resistor, so that enough current passes through the lower triode to give satisfactory g_m , but only a small proportion passes through the upper triode; thus the volts dropped across the load are not excessive. Since the cascode gain depends on the product of the load resistance and the g_m of the lower valve, high gain and therefore excellent stabilization result. Attree's circuit in one form appears in *Figure 37.8*. The 220k resistor is the one responsible for the excellent performance.

When a stabilized negative as well as positive supply is required, two separate stabilizing systems, each with its own transformer and rectifier, may be used, one with its positive terminal and one with its negative terminal earthed, as shown in *Figure 37.9*. In general two mains transformers are necessary (since components possessing two HT secondaries are not stock items) and two voltage reference sources, giving a somewhat expensive arrangement. Using series stabilization it is not possible to derive both positive and negative supplies from a single HT secondary winding, but it can be done using shunt stabilization; the circuit of *Figure 37.10* was devised by K. E. Machin and, besides having only one HT secondary, only employs a single voltage reference source as well. The mode of operation is as follows:

Whereas the action of the series stabilizer is to absorb fluctuations in supply voltage across the series valve, leaving load voltage constant, the shunt stabilizer works by absorbing fluctuations in the load current, so that a

HT SUPPLIES

constant load is presented to the rectifier circuit. Parenthetically it may be remarked that the shunt stabilizer wastes the least power when the device is fully loaded, whereas the series stabilizer is least wasteful when it is lightly loaded.

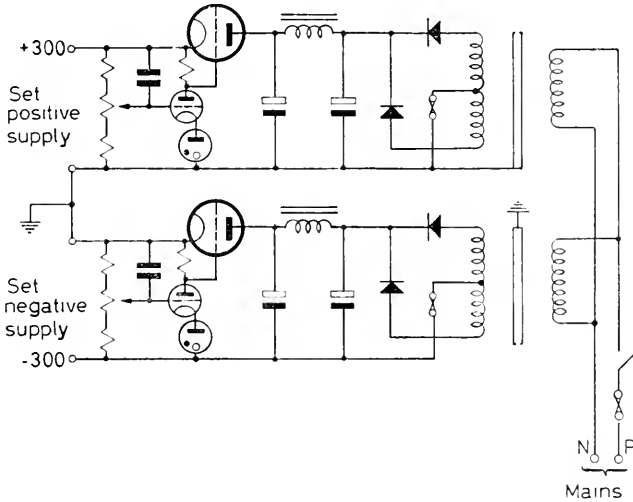


Figure 37.9

The steady current drain on the negative supply produces a sufficient drop across the 5k resistor to ignite the voltage reference tube, applying reference potential to the left-hand side of the differential amplifier, between

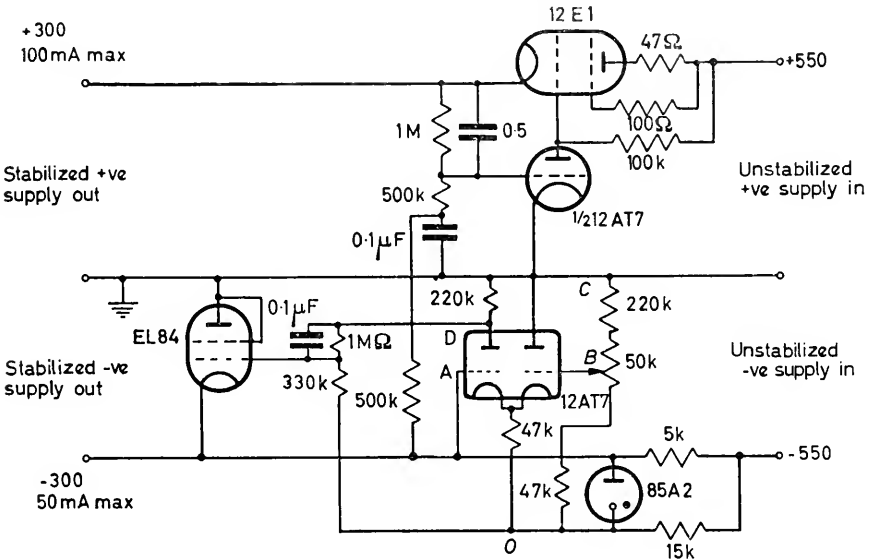


Figure 37.10

A and *O*. The output voltage plus the reference voltage* appears between *O* and *C*, and a definite fraction between *O* and *B* is applied to the other input of the differential amplifier. Any error in the output voltage produces a signal at *D* which is fed forward to the absorber valve grid. The sense of the connections is such that a tendency for the negative output voltage to increase, increases the absorber valve current and increases the voltage drop produced across the 5k resistor.

The upper part of the circuit, concerned with the positive supply, uses conventional series stabilization, but notice how the input to the amplifier valve is derived by comparing the positive output with the negative output rather than with a second reference tube, thus achieving a useful saving in components.

Wide-range HT supplies

When developing a new piece of apparatus it is extremely useful to have a power supply yielding HT voltages continuously variable over a wide range, say 0 to 350 V. Control by variable series resistor is not satisfactory because it gives very poor regulation. Potentiometric control is not very practicable unless only a very light load is to be drawn, since for the output

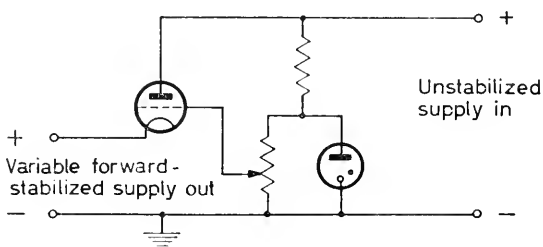


Figure 37.11

impedance to be small at all settings of the output control a value of potentiometer is required so low that it bleeds a great deal of current. What we can do, however, is to supply the power via a cathode follower whose grid potential is variable. Since the grid current is negligible the grid may be potentiometrically fed from a component of high resistance, and therefore, since the potentiometer current is low, it may easily be cold-cathode-stabilized. The scheme is shown in outline in *Figure 37.11*. Notice that the output is described as 'forward-stabilized'. It is protected from mains voltage fluctuations but not from load current fluctuations, since the output impedance is not very low, being $1/g_m$ for the cathode follower, a few hundred ohms. Nevertheless such a device is extremely useful, and a version used by the author appears in *Figure 37.12*.

It is possible to combine the functions of control and rectification in a single valve. When this is done it is called a cathode-follower-rectifier. A power supply along these lines has been described by Walker⁵ (≈ 400 V, ≈ 100 mA) and another by Perry⁶ (≈ 350 V, 40 mA).

* The addition of the reference voltage is incidental to the operation. It allows a larger fraction of the output to be used for control purposes.

HT SUPPLIES

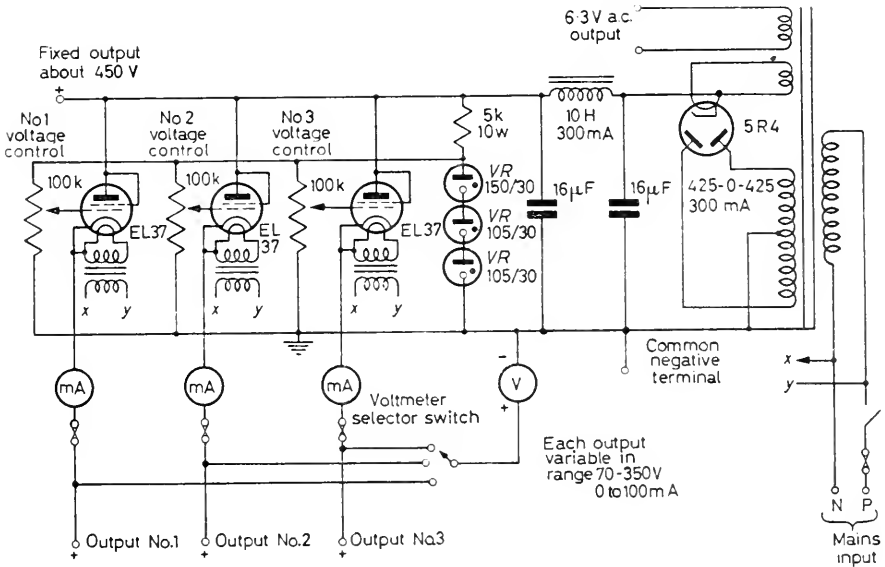


Figure 37.12

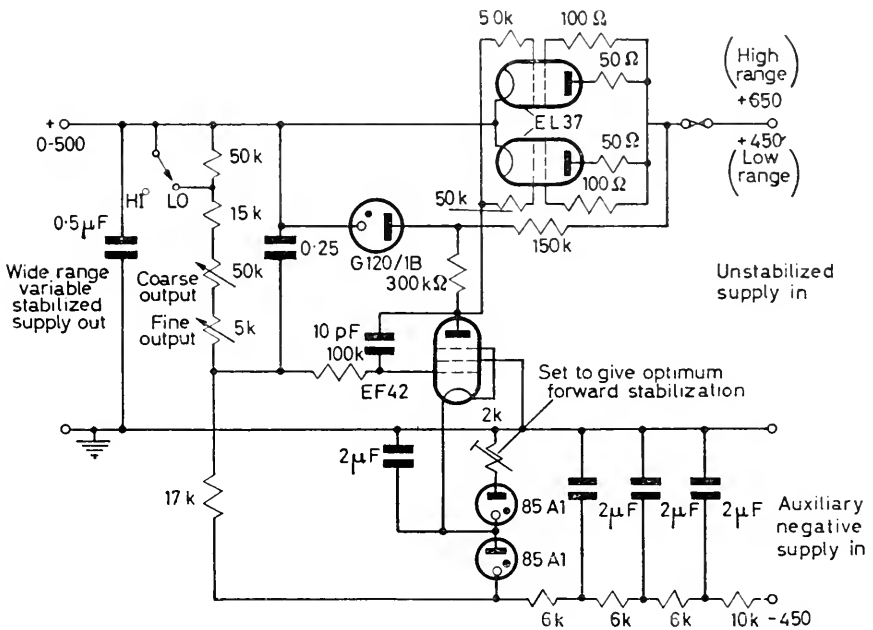


Figure 37.13

POWER PACKS

To produce electronically forward- and backward-stabilized power supplies, having output impedances of a fraction of an ohm, which are also widely variable in voltage, requires a fairly elaborate circuit. The difficulty with the conventional electronically stabilized supply is that the output voltage cannot be reduced to zero because the voltage amplifier valve—on whose action the device depends—is fed from it. The difficulty may be overcome by feeding the cathode of the amplifier from an auxiliary stabilized negative supply, so that volts are still available for the valve even when its anode potential is zero with respect to earth. A stabilizer, due to Scroggie⁷, working on these principles is shown in *Figure 37.13*.

HEATER POWER SUPPLIES

For run-of-the-mill applications such as heater power for pulse circuits and high level amplifier stages, a.c. excitation from a winding on the mains transformer is perfectly satisfactory (*Figure 37.14*). It is usual to earth one

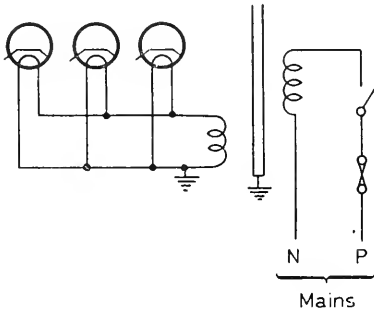


Figure 37.14

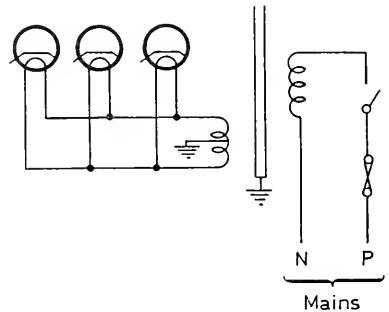


Figure 37.15

side of the winding in order to define the absolute potential of the heater circuit. The only point to watch is that, in so doing, the heater-cathode potential difference rating for none of the valves becomes excessive. For most valves this rating is of the order of 100 V and as, in most circuits, the cathodes operate at a potential not far from earth there is no problem here. Nevertheless there are exceptions, of which probably the most important is the thermionic diode, which often appears in circuit applications where the cathode is very considerably positive or negative to earth. When this happens it may be necessary to supply the diode heater from a separate transformer winding. The heater-cathode rating for valves can be found in the makers' literature.

When low-level amplifier valve heaters are fed from raw a.c. there is a risk that 50-cycle 'hum' may be introduced into signal circuits, either by capacitive or inductive coupling. The effect of capacitive coupling may be reduced by 'balancing' the heater circuit about earth so that the disturbing effects of the two heater leads cancel. One procedure is to provide the heater winding with a centre-tap and earth this (*Figure 37.15*) but the method does not allow for the effects of physical asymmetries in the wiring.

HEATER POWER SUPPLIES

A better method is to use a 'hum-dinger.' This is a low resistance potentiometer—some 50 ohms—connected as in *Figure 37.16*. The slider is then adjusted until the 'hum' is minimized. The effect of inductive coupling is reduced by tightly twisting the heater wiring, so that the magnetic fields due to the two wires as nearly as possible cancel (*Figure 37.17*) and by the

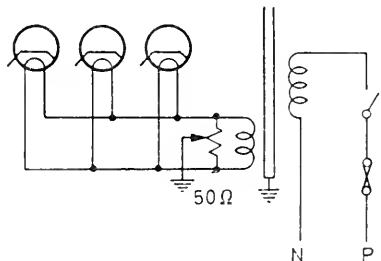


Figure 37.16

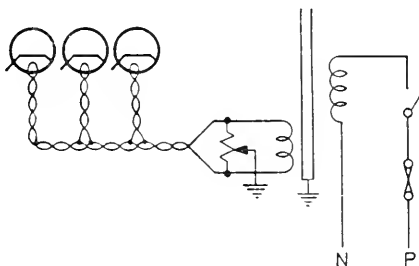


Figure 37.17

use of valves having heaters of special construction aimed at minimizing magnetic field production, such as the Mullard EF86, Osram Z729, etc.

In the highest class of amplifier it may be safer to excite the heaters of low-level stages from direct current, and if the apparatus is direct coupled, a stabilized heater supply may be advisable as well. The simplest d.c. heater supply which also is fairly stable is the lead-acid accumulator, but the cells should be large—i.e. a 6 V car battery—so that the method is bulky and, in so far as accumulators need careful maintenance, not very convenient. It is possible to make heater supplies by transformation,

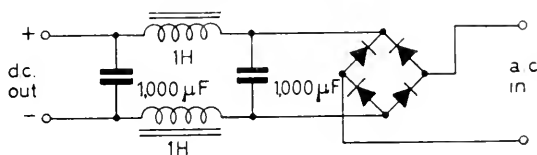


Figure 37.18

rectification and smoothing of the mains, but here again a bulky apparatus results because adequate smoothing of large currents calls for chokes which are large (because the wire must be thick) and capacitors which are large (because of the magnitude of current they have to supply between rectifier conduction periods without significant fall in potential difference). Thus Dickinson⁸ suggests, in a d.c. heater supply of 1 to 2 amps, the use of a 1 H choke in each lead and 1,000 μF input and smoothing capacitors (*Figure 37.18*).

In general the heater current demands of apparatus is constant so that in order to stabilize such a supply it is only necessary to stabilize against variations in mains voltage. A simple device for doing this, which gives a modest but useful output, is due to Cherry and Wild⁹. Stabilizer tubes are used in a bridge circuit on the primary side of the transformer (*Figure 37.19*). With this device a 10 per cent change of mains voltage produces only a 0.35 per cent change of output voltage.

POWER PACKS

More output and better regulation may be obtained with a device described by Attree¹⁰ which is primarily a stabilizer having an a.c. output, but which can easily be arranged to give d.c. This uses two beam tetrodes arranged as a screen-coupled multi-vibrator to generate square waves at 2 kc/s. These are transformed to the voltage required for the output, and are also used to

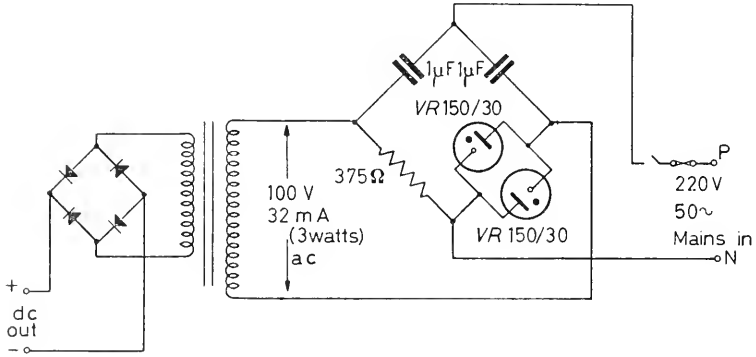


Figure 37.19

heat the cathode of a saturated diode, the anode current of which controls, via a valve, the power supplied to the multi-vibrator (Figure 37.20). In this manner the output R.M.S. voltage is both back-stabilized (output impedance 0.03 ohms for 6 V output) and forward-stabilized (0.1 per cent change in output for 10 per cent change in mains): the a.c. power output available is 10 W. Because of the high frequency and the square waveform this

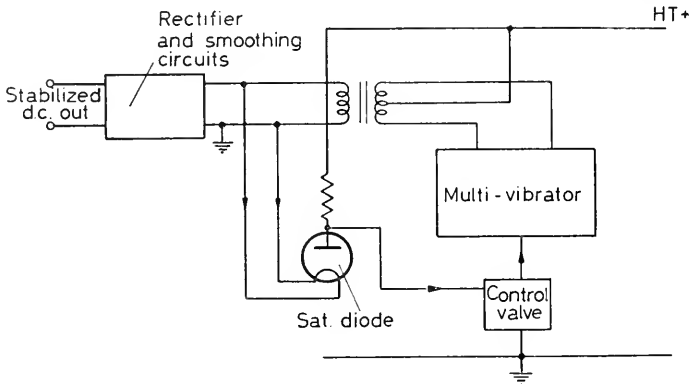


Figure 37.20

supply is particularly easy to rectify and smooth; Attree suggests the use of a single 100 mH choke (actually, the secondary winding of a small loud-speaker transformer) and capacitors of 50 and 1,000 μF . At 2 A output the ripple is then below a millivolt; unfortunately 3 W are lost in the rectification and smoothing process.

A recent device of promise is the electrolytic smoothing and stabilizing cell, manufactured by L'Accumulateur Étanche S. A. of Brussels. A heater

EHT SUPPLIES

supply derived by rectification of the a.c. mains may be smoothed by floating an ordinary accumulator across it, but if this is done care has to be taken to arrange matters so that the accumulator is neither gradually discharged—because the rectified voltage is too low—nor the water all electrolysed off by overcharging because the rectified voltage is too high: in addition the usual maintenance procedures are necessary. In these special smoothing and stabilizing cells the chemical cycle is closed, no gas is evolved and the device is hermetically sealed. Any current up to a stated maximum can be passed through the cell and overcharging is impossible; therefore the external circuit can be arranged to pass a steady 'charging' current through, and the cell will maintain a steady potential difference across itself of between 1.4 and 1.45 V. The temperature coefficient is about -0.17 per cent/ $^{\circ}\text{C}$ and the maximum current ratings lie between 20 mA for a model weighing 0.1 oz., and 1 amp for one weighing $2\frac{1}{2}$ oz. The effective equivalent capacitance of the 400 mA version is stated to be 60,000 μF . Further details may be had from Mercia Enterprises, Ltd., 30 Silver Street, Coventry.

EHT SUPPLIES

For *ad hoc* apparatus such as, for example, might be used in vision experiments involving photomultiplier cells, it may be worth considering deriving EHT from a chain of series-connected dry batteries of the layer type. Such a supply has the merits of stability and freedom from noise—provided the batteries are reasonably new—and may easily be earthed at either pole or at any intermediate point as required. For more permanent installations, however, mains-driven supplies are more satisfactory and fall into two types, those in which EHT is generated directly by transformation, rectification and smoothing of the mains, and those which involve the generation of radio frequency oscillations.

Direct methods

Because the current demands on EHT supplies are low—seldom more than a milliamp—two useful simplifications follow: half-wave, rather than full-wave, rectification is adequate, and resistance, rather than choke, smoothing is sufficient. A half-wave positive supply may have the configuration of *Figure 37.21a* or *b*. Similarly a negative supply could be as in *Figure 37.21c* or *d*. However, it is not hard to see that, whereas in *a* and *c* the maximum voltage the transformer insulation has to stand is equal to the peak secondary voltage, in *b* and *d* the maximum voltage stress is nearly twice as much. For this reason the former configurations are general, and EHT transformers usually have one end of the secondary winding internally earthed. It follows that when a thermionic rectifier valve is used (*Figure 37.22*) the transformer has in addition to have a special rectifier heater winding, whose insulation also must be able to stand the peak secondary voltage to earth, and in the case of the positive supply must in addition withstand twice the peak voltage between itself and the live end of the main secondary.

From this it is clear that non-thermionic rectifiers are advantageous, quite apart from their longer life, in that—particularly for positive supplies—they make possible the use of a cheaper transformer. As EHT transformers are

POWER PACKS

expensive—it is poor economy not to buy the best—any saving here is welcome. Smoothing and input capacitors are commonly between 0.1 and 0.5 μF and the smoothing resistor about 100k.

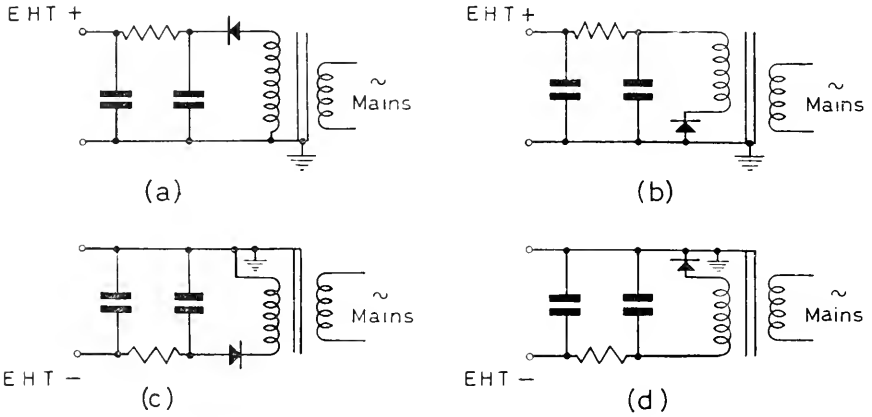


Figure 37.21

There are two important objections to the direct generation of EHT. One is that a capacitor of, say, 0.25 μF rated at 5,000 V is large and expensive. The other is that, when charged, it is extremely dangerous. Radio-frequency EHT methods overcome both these difficulties.

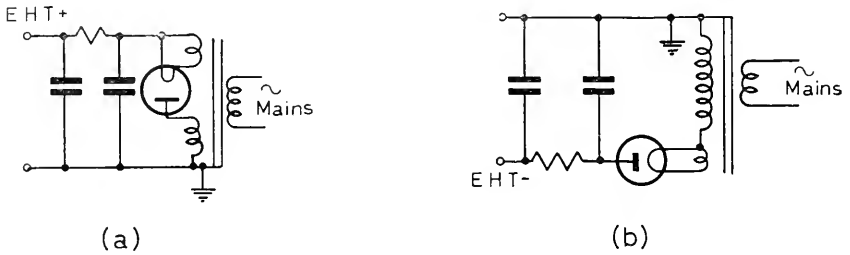


Figure 37.22

RF EHT

In this method power is drawn from a HT supply to feed a radio-frequency LC oscillator. The oscillator inductance forms the primary of a step-up transformer, whose secondary voltage is rectified and smoothed to form the EHT. Because the frequency is high the necessary high voltage smoothing and input capacitors can be quite small, of a capacity insufficient to hold a lethal charge. The step-up transformer is a simple air-cored affair, and the overall cost, despite the extra valve and associated small components, is probably no greater than that incurred in generating EHT from 50 cycles.

An RF supply giving $-2,000\text{ V}$ at 1 mA has been described, including details for making the RF transformer, by Dickinson¹¹ and has the circuit of Figure 37.23. Notice that the EHT rectifier circuit is of the type *d* in Figure 37.21. This throws more strain than is absolutely necessary on the transformer insulation, but allows Dickinson to heat his rectifier valve from the common heater supply.

EHT SUPPLIES

The design of RF EHT transformers has been discussed by Barron¹². Alternatively a range of ready-made coils are available from a number of firms specializing in EHT generation equipment.

Stabilization of EHT is particularly easy to accomplish using the RF method. Lowe¹³ describes a circuit for supplying Geiger-Müller tubes

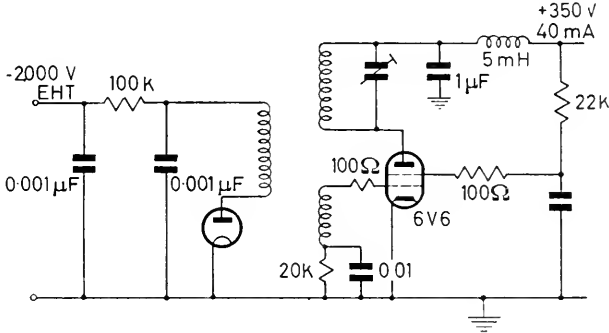


Figure 37.23

which has a change of output voltage of less than 0.1 per cent for a 10 per cent change in the HT voltage supplying the oscillator, or for a variation in output current from zero to 32 μ A (full load). The scheme is shown in outline in Figure 37.24. A fixed fraction of the output voltage is taken by R_1 and R_2 and applied to the direct-coupled amplifier valve V_1 . Since the

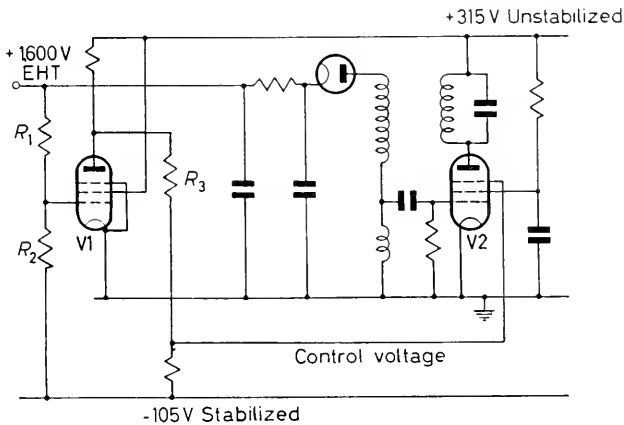


Figure 37.24

—105 V line is stabilized, V_1 grid potential has a definite value when the EHT voltage is correct. Departures of the EHT from the desired value are amplified in V_1 and transferred via R_3 to the suppressor of the RF oscillator valve V_2 in such a sense as to correct the departure. The oscillator and rectifier circuits are closely related to Dickinson's in Figure 37.21a, except that Lowe uses Figure 37.21a type rectification and heats his rectifier cathode from a separate RF oscillator, not shown here.

STABILIZED AC MAINS SUPPLIES

As an alternative to stabilizing separately EHT, HT and heater supplies it may be worth considering stabilizing the supply mains itself; that is, to draw all one's a.c. power supplies from a single stabilizing device which is fed with unstabilized National Grid power. Stabilized a.c. is in any case useful for supplying experimental apparatus of a non-electronic nature, such as light sources. There are a number of commercial units which do this, and designs have been published from time to time which are suitable for making up. There are various methods of approach, but all—apart from the constant voltage transformer—are the same in principle: a sensing device detects discrepancies between the actual output voltage and that required, and causes a regulator to correct them—at least partially. Thus in a device due to Ackland¹⁴ the output voltage is bridge-rectified and compared with that existing across a voltage reference tube. The difference is used to actuate a sensitive relay, which—via other relays—controls a motor-driven variac transformer. Such a device is clearly capable of stabilizing large powers, but the rate of response to a sudden change in output would not be very rapid; Ackland claims a correction rate of 8 V per second. A much faster type of a.c. stabilizer is due to Benson and Seaman¹⁵. Here a saturated diode is the sensing unit; the diode output is amplified by a valve whose anode current is used to vary the saturation of part of the core of an auto transformer, so that its effective step-up ratio is altered. The response time is $\frac{1}{4}$ second.

STABILIZED-CURRENT POWER SUPPLIES

It happens occasionally that a constant current, rather than constant voltage, power supply is required; for example, in the production of a stable magnetic field by a coil whose temperature, and hence resistance, is subject to

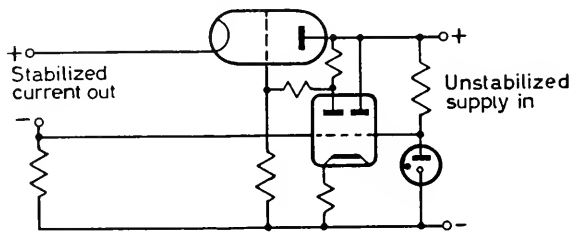


Figure 37.25

fluctuation (the author has in mind the tachometer-generator field winding of a velodyne).

A constant current characteristic may be had by electronic stabilizing techniques very similar to those used for producing a constant voltage; instead of a fraction of the actual output *voltage* being fed to the differential amplifier, a voltage is derived proportional to the output *current* by passing the latter through a high stability resistor (Figure 37.25). The performance of circuits of this type is discussed by Sowerby¹⁶, who shows that in a typical

REFERENCES

case the effective output impedance of the device is 1·2 megohms. Clearly a 1 per cent change in resistance of a 100 ohm coil in series with 1·2 megohms has a quite negligible effect on the coil current.

'Stopper' resistors

In many practical circuits, particularly those containing a power valve, resistors connected immediately in series with the control grid, screen grid or anode sometimes appear. These were not discussed in Part I.

The wiring associated with the valve has stray capacitance and a small amount of inductance. Under certain circumstances an unwanted LC oscillator is formed. These 'parasitic' oscillations are extremely rapid, of the order of 100 Mc/s and may be difficult to detect as such; but the normal operation of the circuit is impaired. To ensure they cannot occur it is a sound precaution to damp heavily any possible such spurious LC circuits by the inclusion of 'stopper' resistors. Anode and screen stoppers may be of the order of 100 ohms, and control grid stoppers 10 k; they should be soldered in as close as possible to the relevant valveholder tag.

Stoppers appear in *Figures 37.8, 37.10, 37.13, 37.23*, etc.

REFERENCES

- ¹ LUKE, R. J. *Electronic Engineering* 28 (1956) 100
- ² *Philips Technical Review*, February 1941
- ³ VALLEY, G. and WALLMAN, H. *Vacuum Tube Amplifiers*. Radiation Lab. series, McGraw-Hill
- ⁴ ATTREE, V. H. *Electronic Engineering* 27 (1955) 174
- ⁵ WALKER, A. H. B. *Wireless World* 58 (1952) 374
- ⁶ PERRY, B. J. *Electronic Engineering* 28 (1956) 517
- ⁷ SCROGGIE, M. G. *Wireless World* 54 (1948) 543
- ⁸ DICKINSON, C. J. *Electrophysiological Technique*, *Electronic Engineering*
- ⁹ CHERRY, L. B. and WILD, R. F. *Proc. Inst. Radio Engrs. N.Y.* 33 (1945) 262
- ¹⁰ ATTREE, V. H. *Electronic Engineering* 25 (1953) 71
- ¹¹ DICKINSON, C. J. *Wireless World* 57 (1951) 70
- ¹² BARRON, J. *Electronic Engineering* 26 (1954) 393
- ¹³ LOWE, A. E. *Electronic Engineering* 27 (1955) 85
- ¹⁴ ACKLAND, R. G. *Electronic Engineering* 26 (1954) 143
- ¹⁵ BENSON, F. A. and SEAMAN, M. S. *Electronic Engineering* 28 (1956) 260
- ¹⁶ SOWERBY, J. MCG. *Wireless World* 55 (1949) 395

STIMULATORS

The classical device for stimulating biological tissue is, of course, the induction coil. By interrupting the direct current flowing in a primary winding an e.m.f. is induced into a mutually coupled secondary which drives stimulating current through the preparation* (*Figure 38.1*). The stimulating current

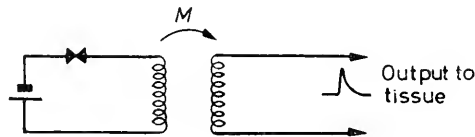


Figure 38.1

waveform has the form shown; a sharp rising phase followed by an exponential decay. The magnitude of the shock depends on the mutual inductance between the coils—one is arranged to slide in and out of the other—and the time constant of the decay depends on the self inductance of the secondary winding and the resistance of the preparation, being equal to the

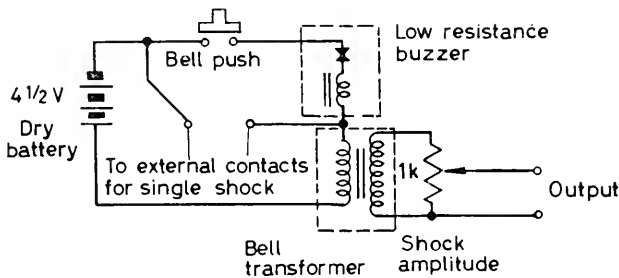


Figure 38.2 The bell transformer is used with the windings reversed; that is, the winding intended for connection to the mains is used as the secondary

former divided by the latter. Two modes of operation are possible; repetitive stimulation, in which the contacts open and close automatically in the same way as in an electric bell, and 'single shock', in which external contacts are opened once only, at a definite instant, usually by the rotation of a kymograph.

Induction coil stimulators are expensive to buy and tedious to make. Fortunately it is easy and cheap to make up a small unit which fulfils a similar purpose from readily obtainable components as shown in *Figure 38.2*. The operation is self-explanatory. At Cambridge these small stimulators are

* A 'make' shock is also delivered when the primary circuit is closed; but since the amplitude of the shock depends upon the rate of change of primary current, and since the rate of change of primary current upon closing the circuit is limited by the self inductance of the primary winding, the 'make' shock is much smaller.

STIMULATORS

assembled in tin boxes measuring only $7 \times 4 \times 2\frac{1}{4}$ in. complete with battery, and have proved satisfactory for medical students' practical work.

For research work something rather more elaborate is required. In repetitive operation the repetition frequency, in addition to the shock amplitude, ought to be controllable over a wide range. A well-trying way of achieving this is the simple 'neon stimulator' (Figure 38.3) which is the relaxation

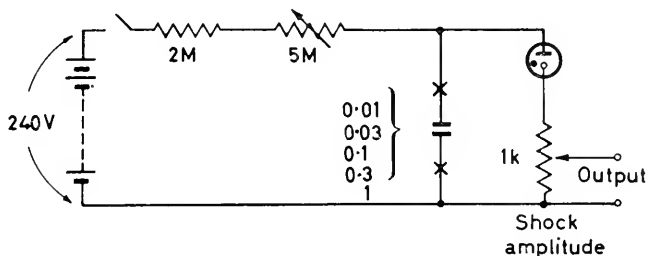


Figure 38.3 The symbol $-\times \cdots \times-$ means that the component between the crosses is switchable to any of the values shown

oscillator described in Chapter 7. Coarse control of shock frequency is obtained by switching in various capacitors between, say, 0.01 and $1 \mu\text{F}$, and fine control by having part of the charging resistance variable. For the 'neon' lamp a difference diode characteristic is desirable, but in point of fact almost any kind of cold-cathode diode will do. Old voltage stabilizer tubes whose regulating properties have long since departed are quite suitable; sometimes more output is obtainable if they are connected up the wrong way round.

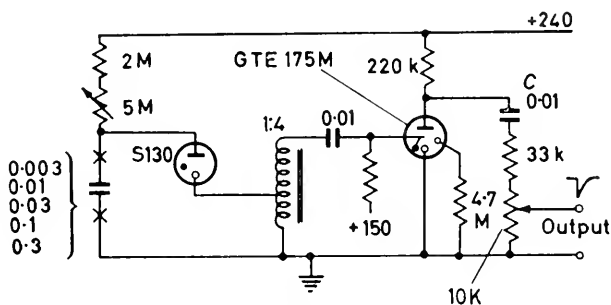


Figure 38.4

A drawback of the simple neon stimulator is that the time constant, and to some extent the amplitude, of the output depends on the charging capacitor in use; that is, frequency and shock waveform are not completely independent. If this is undesirable recourse may be had to two cold-cathode valves, one to generate the repetitive frequency and one to fashion the shock. The circuit of Figure 38.4, devised by the author, gives shocks of from zero to 10 V amplitude at any repetition frequency between 1 and 100 per second. The circuit associated with the S130 tube is a straightforward relaxation oscillator, like Figure 38.3, except that the load is not a potentiometer but

STIMULATORS

the primary of a 1 to 4 step-up transformer. The cold-cathode trigger tetrode has its HT arranged so that the main gap just does not break down, and the 0.01 capacitor C charges to 240 V. Each time a pulse arrives from the oscillator via the trigger electrode, the tube fires and the capacitor discharges via the 10k potentiometer across which the output is developed. The 33k resistor is to limit the discharge current. On the completion of discharge the capacitor rapidly re-charges to 240 V via the 22k resistor. The waveform of the output pulse is now independent of the frequency settings.

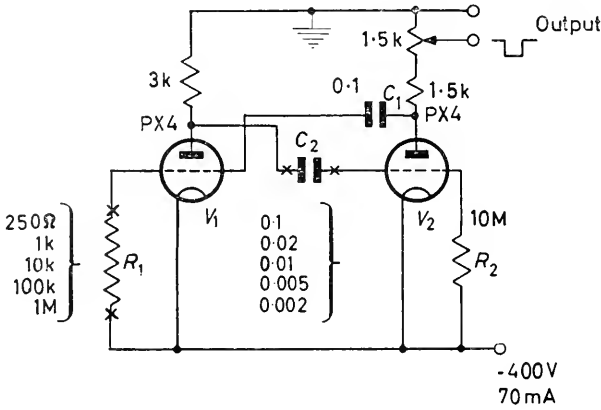


Figure 38.5

The magnitude of the output of 'neon' or cold cathode stimulators is of the order of 10 V, sufficient for excised preparations but of little use where there is much by-passing of stimulating current by surrounding structures. More powerful outputs can be had from thyratrons¹, but there seems to be an increasing tendency among research workers to demand stimulating pulses of rectangular waveform; the repetition rate, duration and amplitude all to be variable. To do this it is usual to employ hard valves. An early hard-valve stimulator was described in 1944 by Ritchie²; this comprised a pair of power triodes arranged as a multi-vibrator (Figure 38.5). The output is delivered when V_2 conducts and has duration determined by the time constant C_1R_1 . The range of pulse durations obtainable was 25 μ sec to 100 msec. The interval between pulses, when V_1 conducts and V_2 is cut off, is determined by R_2C_2 and ranged from 1 to 1/50 of a second. Notice that when the output is derived from an anode load the *positive* power supply terminal is earthed.

A drawback of a simple stimulator such as this is that when the shock duration and shock interval are comparable, and the duration is varied, the shock frequency is varied too. On the whole it is probably more satisfactory to be able to set a definite frequency, independent of the shock duration, and this may be done by using a multi-vibrator to control a flip-flop. Two rather similar stimulators along these lines are due to Bernstein³ and to Ead⁴. Both of these have in addition an output stage to buffer the flip-flop from the effect of variations in the load. Bernstein uses an output valve arranged as an amplifier and derives a negative-going stimulus pulse. Ead

STIMULATORS

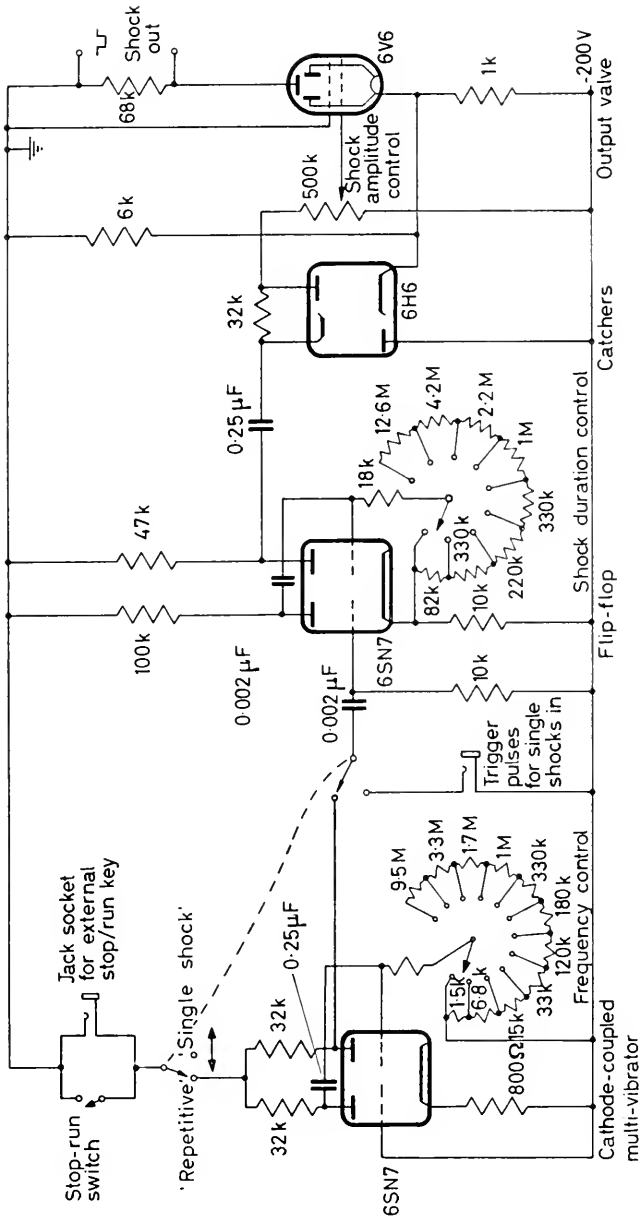


Figure 38.6

STIMULATORS

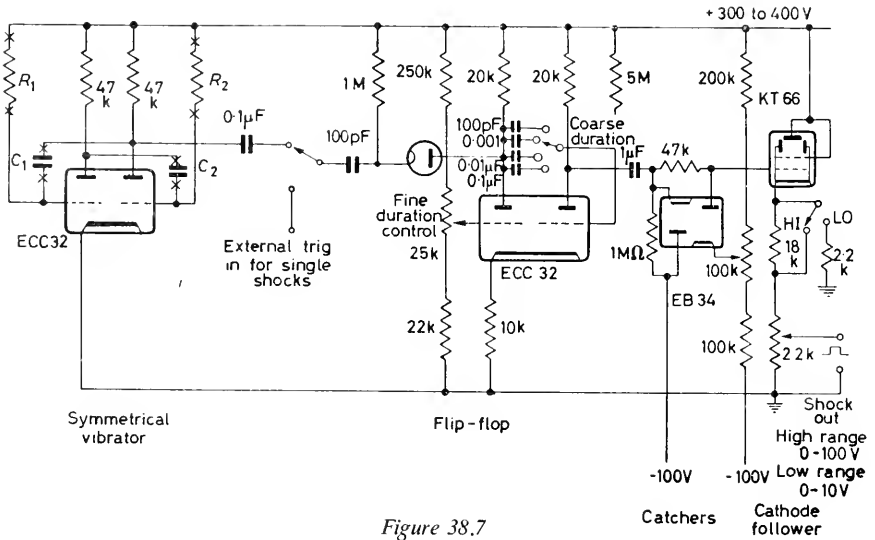


Figure 38.7

Table of Switched RC Values in Ead's Stimulator

Repetition rate (shocks/second)	$R_1(M\Omega)$	$C_1(\mu F)$	$R_2(M\Omega)$	$C_2(\mu F)$
1/20	20	0.05	20	2
1/10	20	0.05	10	2
1/5	20	0.05	4.7	2
1/2	20	0.05	1.5	2
1	20	0.05	20	0.05
2	20	0.001	20	0.05
5	20	0.001	10	0.05
10	20	0.001	4.7	0.05
20	20	0.001	1.5	0.05
50	20	0.001	20	0.001
100	10	0.0001	20	0.001
200	10	0.0001	10	0.001
500	10	0.0001	3	0.001
1,000	10	0.0001	1	0.001

COUPLING THE PREPARATION TO THE STIMULATOR

uses a cathode follower to produce a positive-going output. Both stimulators have facilities for disconnecting the flip-flop from the multi-vibrator and driving it instead from an external pulse so that single-shock operation is obtained. Bernstein's is shown in *Figure 38.6*; it has repetition frequencies from 0.2 to 1,000 per sec and pulse durations from 100 μ sec to 50 msec. Ead's has frequencies from 0.05 to 1,000 per sec and durations from 10 μ sec to 100 msec; it is shown in *Figure 38.7*. Notice in these circuits the use of catching diodes to define the limits of grid potential applied to the output valves.

In recent years a rather more elaborate pattern of stimulator has emerged, having the block diagram of *Figure 38.8*. At each stroke of the oscillator

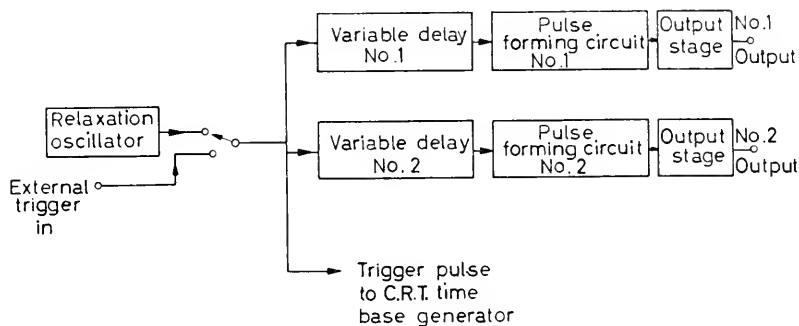


Figure 38.8

the-time base sweep on the cathode ray tube is initiated, and one or two shocks are delivered to the preparation as required, after a delay. Apart from the value in research of being able to investigate the effect of two stimuli delivered at different times, the use of delays ensures that nothing of physiological interest can happen until the time base is 'on its way', and there is no risk of phenomena being missed at the beginning of the trace. The first stimulator of this kind was described by Attree⁵ and another was published later by Perkins⁶. Attree uses a Miller transitron oscillator, a transitron delay circuit, a pentode flip-flop pulse-forming circuit and a cathode follower output. In addition there are a number of other refinements. Perkins uses a multi-vibrator, a flip-flop delay circuit, flip-flop pulse-forming circuit and cathode follower output. The author has also designed a stimulator of this type, which is shown in *Figure 38.9*, but before discussing it, it is necessary to digress into an important problem of technique: how best to couple the preparation to the stimulator.

COUPLING THE PREPARATION TO THE STIMULATOR

If one is using electrical stimulation and, say, observing the rate of secretion of a gland or the contraction of a muscle, no problem arises. The output connections are applied to the preparation via electrodes as dictated by the experiment. The difficulty arises when one is using electrical recording as well, where stimulus pulses of the order of volts are being applied to the preparation, and perhaps, only a few cm away, recording electrodes—which,

STIMULATORS

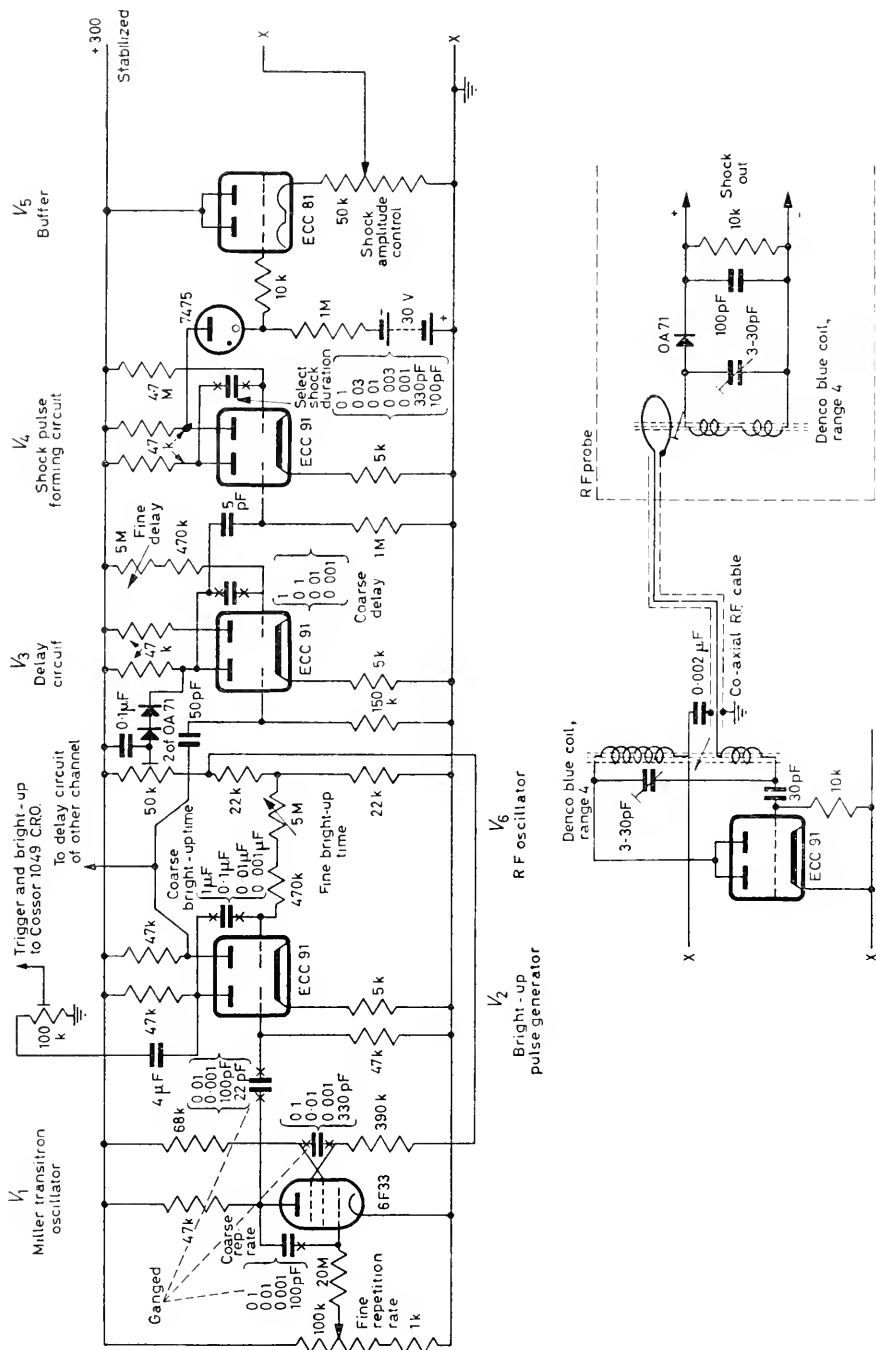


Figure 38.9

COUPLING THE PREPARATION TO THE STIMULATOR

by their very nature cannot be screened—are measuring signals perhaps of the order of microvolts. On stimulating the preparation the record may be completely obscured by *stimulus artefact*, i.e. direct conductive and capacitive coupling between stimulating and recording electrodes. Furthermore, one

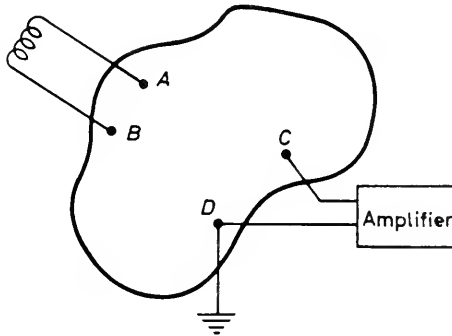


Figure 38.10

may make the distressing discovery that the preparation is stimulated with only one stimulating electrode in place, or even that the site of stimulation is nowhere near the stimulating electrode at all!

The situation is represented rather artificially in Figure 38.10. The irregular lump is the preparation, A and B are stimulating electrodes, C is a recording electrode feeding a single ended—for argument's sake—recording amplifier,

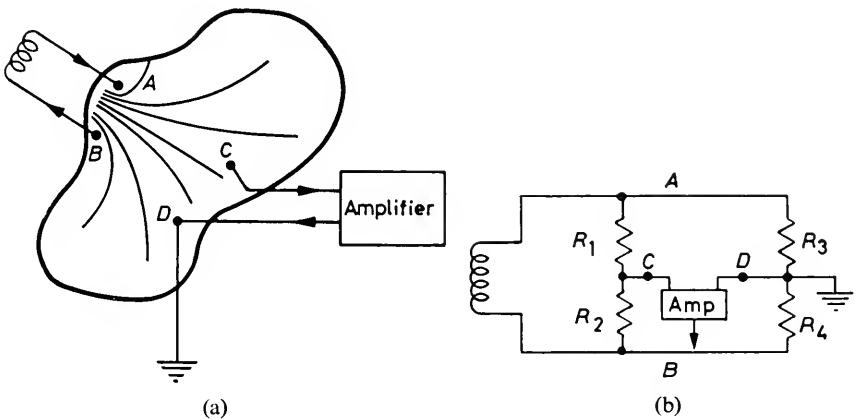


Figure 38.11

and the preparation is earthed at D. The stimulator is represented in the traditional manner as the secondary of an induction coil.

Now let the stimulator pass a steady current. A system of equipotential surfaces will be set up in the body of the preparation in some such manner as Figure 38.11a, and it is clear that C and D will in general assume different potentials, both intermediate between A and B. We can represent this state

STIMULATORS

of affairs by two potential dividers, R_1R_2 and R_3R_4 , as shown in *Figure 38.11b*. We now put in the stray capacitances between *A* and *B* and the recording electrode, calling them C_1 and C_2 , and between *A* and *B* and earth, calling them C_3 and C_4 , giving us *Figure 38.12a*. Finally, let the stimulator now become a constant-voltage pulse generator, of output E , and let the amplifier be replaced by its input impedance, Z , assumed to be relatively large (*Figure 38.12b*).

The problem of minimization of stimulus artefact is now seen to be that of minimizing the out-of-balance current in a bridge circuit, but before

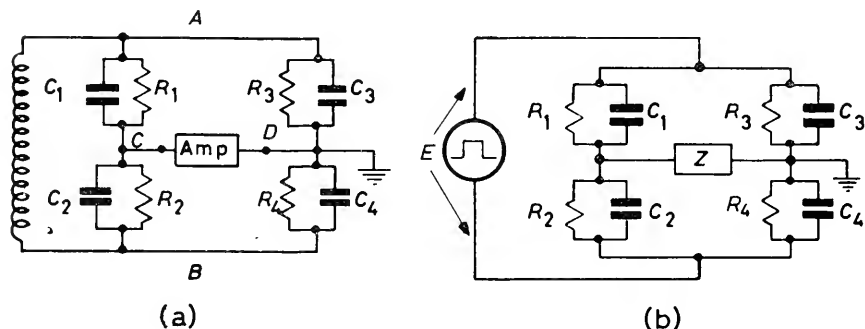


Figure 38.12

attending to this it is worth remarking the nature of the unbalance current. It has three components: a transient of magnitude

$$E \left\{ \frac{C_2}{C_1 + C_2} - \frac{C_4}{C_3 + C_4} \right\}$$

a final value which it reaches if the pulse goes on long enough of magnitude

$$E \left\{ \frac{R_1}{R_1 + R_2} - \frac{R_3}{R_3 + R_4} \right\}$$

and a phase between them which is the difference of two exponentials corresponding to the time constants

$$(C_1 + C_2) \left(\frac{R_1 R_2}{R_1 + R_2} \right)$$

and

$$(C_3 + C_4) \left(\frac{R_3 R_4}{R_3 + R_4} \right)$$

The observed artefact on the cathode ray tube face may have any of the forms in *Figure 38.13*.—but only *a*, *b*, or *c* are at all likely. It is also worth noting at this point that the bridge cannot possibly be balanced if one side of the stimulator output is earthed, i.e. if R_3 or R_4 is zero. This puts all the

COUPLING THE PREPARATION TO THE STIMULATOR

electronic stimulators we have discussed so far—as they stand—straight out of court.

To reduce the artefact one can either:

(1) attempt to balance the bridge by juggling with the ratios of the impedances of the arms.

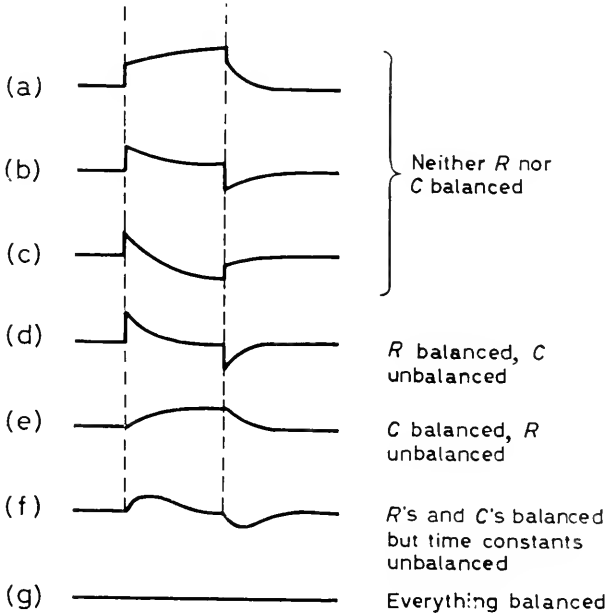


Figure 38.13

(2) make the impedances of all the arms as high as possible. Applying Thévenin's theorem to the bridge, we see that looking into it from the amplifier we have its open-circuit e.m.f., determined by its degree of off-balance, in series with its internal impedance, determined by the impedance of its

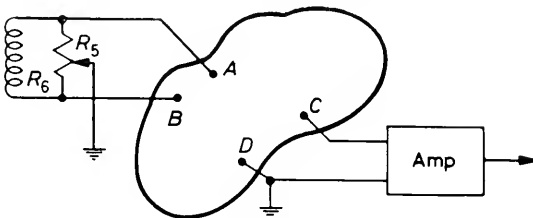


Figure 38.14

arms. If we cannot correct the unbalance current by method 1, we can at least reduce it by keeping the arm impedances as high as possible.

Balancing the bridge—We can partially balance the bridge by what might be described as 'brute force'—using a *Wagner earth*. This is done by connecting between the stimulator terminals a potentiometer of much lower resistance than $R_3 + R_4$, and earthing the slider (Figure 38.14). The bridge

STIMULATORS

circuit now has the appearance of *Figure 38.15*, and clearly if $R_5 + R_6$ is much smaller than $R_3 + R_4$, $R_3 + R_4$ are 'swamped'—i.e. have little effect, and 'R-balance' can be effected; the artefact can be simplified to type (d) in *Figure 38.13*. Dickinson⁷ shows a stimulator possessing this facility. Strictly speaking it should be possible to achieve C 'balance' with a differential variable capacitor, and the time constant balance with an additional resistor,

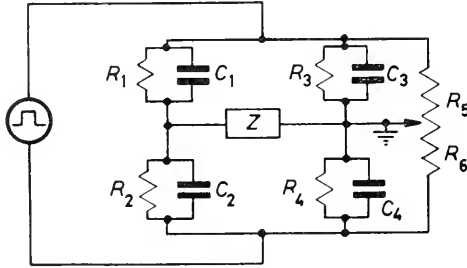


Figure 38.15

in the arrangement shown in *Figure 38.16*, but the author has never heard of anyone attempting it. Probably it would be much too difficult to set up, and in any case Wagner earth-type balancing is a risky procedure. The reason is that the balancing current which flows to earth at point *E* re-enters the preparation at the earthing connection *D*. It is this earth current which is liable to cause the anomalous stimulation mentioned earlier. We therefore fall back on method 2.

Maximizing the bridge arm impedances—In practical terms this means that the site of stimulation should be as far as possible from the recording site

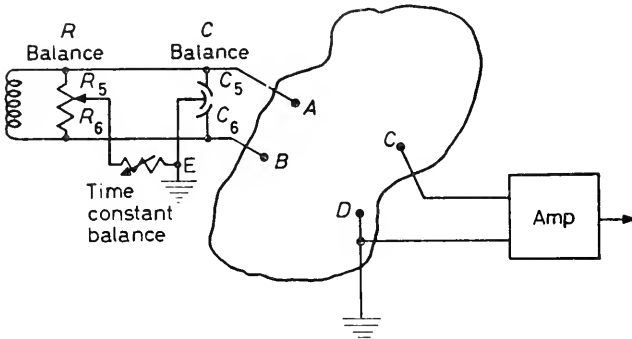


Figure 38.16

(maximum R_1 and R_2) and the stimulator and recording wiring should be separated as widely as possible (minimum C_1 and C_2). So much is common-sense, and this analysis would not be justified if these were the only conclusions. What is less obvious, and equally important, is that the stimulator output circuit should have no conductive connection to earth at all (other than the inevitable path through the preparation—maximum R_3 and R_4) and the minimum possible stray capacitance to earth (minimum C_3 and C_4).

THE RF COUPLED STIMULATOR

One way of achieving this, which is fairly successful, is to pass the stimulator output via an iron-cored transformer of special low-capacitance construction. The snag is that the kind of transformer which has a low capacitance from secondary to primary and from secondary to earth (the core) is precisely the kind of transformer that makes a bad stimulating pulse-transformer; the primary inductance is likely to be insufficient to pass the longer pulses required without sag; and the leakage inductance is liable to be considerable, spoiling the leading and trailing edges. This is the reason for the introduction of the RF coupled stimulator.

THE RF COUPLED STIMULATOR

This ingenious device is due to Schmitt⁸. It consists of a radio-frequency LC oscillator whose coil is inductively loose-coupled to a second coil, tuned to the oscillator frequency. RF voltages appearing across the secondary coil are rectified and smoothed to produce a direct voltage. Since the

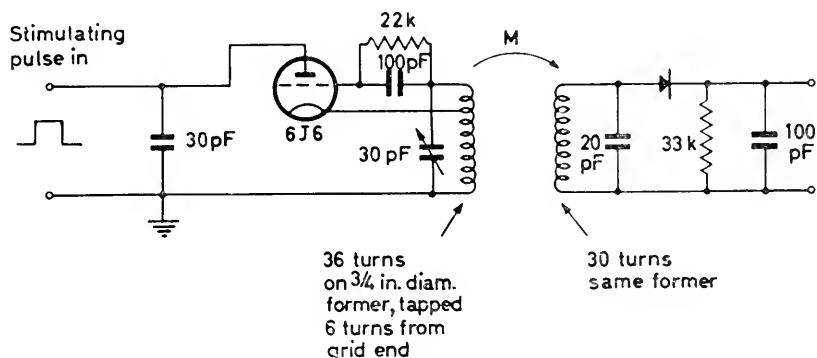


Figure 38.17

amplitude of oscillation is nearly proportional to the HT voltage supplying the oscillator, and since the direct voltage produced by the rectifying and smoothing circuit is nearly proportional to the amplitude of oscillation, it follows that the rectifying-circuit output is a good copy of the oscillator HT. If the stimulator output is used to provide the HT for the oscillator, then the rectifier and smoothed output can be arranged to follow the stimulator output closely, but with this important difference; that by having the secondary circuit 'floating' and only loosely coupled to the oscillator, the direct voltage derived from it appears to come from a generator having a very high impedance to earth; the conditions are fulfilled for avoiding spurious stimulation and minimizing artefact, and the shock can be of unlimited duration. Schmitt's original radio-frequency coupler is shown in Figure 38.17. Other couplers have been described by Haapanen⁹ and Perkins⁶. All these are in the form of 'add-on' units to stimulators having one side of the output earthed. In order to keep the capacitance to earth of the secondary circuit small, the coupler unit must be brought up close to the preparation.

It seemed to the author that the advantages of RF coupled stimulation are so overwhelming that the RF unit might be regarded not so much as an optional luxury attachment but as an essential component. In this case the

STIMULATORS

RF oscillator might well be included within the body of the stimulator so that the coupler unit or 'RF probe' now only contains the secondary circuit and rectifier components; the two are joined together by a single coaxial cable carrying RF power. In this manner a much smaller and neater probe becomes possible.

This has been done in the author's stimulator, one channel of which is shown in *Figure 38.9*. A Miller transitron oscillator feeds flip-flop No. 1, whose negative going square wave output from anode No. 1 is used to brighten and trigger the trace of a Cossor 1049 oscilloscope. The leading edge of the positive going square for anode 2 wave is differentiated to trigger flip-flop 2, which is the delay circuit. At the end of the delay period flip-flop No. 3 generates the requisite stimulus pulse and passes it to a cathode follower.

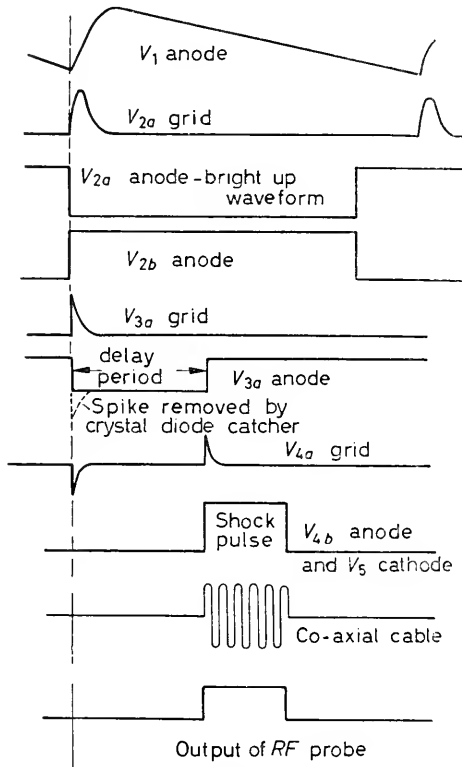


Figure 38.18

A shock amplitude control follows, and transfers the pulse as HT to the oscillator circuit, which is a modified series-fed Hartley. The modification consists in bringing out the oscillator tuned-circuit circulating current through coaxial cable to a single-turn coupling coil in the RF probe. The repetition frequency of this stimulator ranges from 1,000/sec to 1 shock every 35 sec. The maximum delay is 250 msec and the pulse duration is variable between 100 μ sec and 100 msec. The shock amplitude is adjustable in the range 0-40 V. Waveforms are sketched in *Figure 38.18*.

REFERENCES

REFERENCES

- ¹ NASTUK, W. L. and BORISON, H. L. *Rev. sci. Instrum.* 18 (1947) 669
- ² RITCHIE, A. E. *J. sci. Instrum.* 21 (1944) 64
- ³ BERNSTEIN, L. *J. Physiol.* 111 (1950) 34P
- ⁴ EAD, H. W. *J. Physiol.* 113 (1951) 16P
- ⁵ ATTREE, V. H. *J. sci. Instrum.* 27 (1950) 43
- ⁶ PERKINS, W. J. *Electron. Engng.* 27 (1955) 434
- ⁷ DICKINSON, C. J. *Electrophysiological Technique*, Electronic Engineering
- ⁸ SCHMITT, O. H. *Science* 107 (1948) 432
- ⁹ HAAPANEN, L. *Acta physiol. scand.* 29 (1953) 394

BIOLOGICAL AMPLIFIERS

In this chapter we consider types of amplifier for bioelectric potentials. Amplifiers for transducer outputs are dealt with in the appropriate section of Part III.

The amount of gain which can usefully be employed in a recording channel is, as was pointed out in Part I, determined by the noise—including drift—generated by the preparation and the first amplifying stage. Suppose one is looking for spike potentials in a noisy trace on a cathode ray tube. Then the appearance of a spike of amplitude, say, 100 times the R.M.S. noise voltage could be taken to represent a physiological event with a probability amounting to almost complete certainty. If a spike of duration 1 msec and of amplitude 10 times the R.M.S. noise appears in a trace 100 msec in duration, the probability that it represents an action potential rather than a particularly large noise deflection still exceeds 0.99. However, the possibility of spikes smaller than this being noise artefacts has to be taken into account.

In practical terms, an a.c. coupled amplifier with a pass-band extending from 50 to 10,000 cycles might exhibit a noise level of $2 \mu\text{V}$ R.M.S. with the input terminals short-circuited. With the input terminals fed from a glass microelectrode the noise will be very much greater, but in the case of a low resistance preparation and electrode system the noise level should still not be much above $2 \mu\text{V}$ R.M.S. A reasonable procedure would be to arrange that, at full gain, a $20 \mu\text{V}$ event at the input, i.e. one which represents a signal with moderate certainty, produces half of full-scale deflection on the cathode ray tube. That is, it deflects the spot, from the tube centre, half of the way to the top or bottom of the screen. A typical 6 in. cathode ray tube with an EHT of 2 kV has a deflection sensitivity of 0.5 mm/V , so that the voltage required to produce $1\frac{1}{2}$ in. deflection is about 75. The maximum total amplifier gain needed is then $75 \text{ V}/20 \mu\text{V} = 3.75 \times 10^6$, which may be obtained with 4 stages.

It is frequently not possible to use as much as this. In an amplifier habitually to be used with microelectrodes the maximum gain could probably be reduced by a factor of 10. Again, in RC coupled amplifiers of long time constant (e.g. 4 couplings of 4 seconds time constant each) the pass-band would extend down to a frequency range where flicker and carbon resistor noise would become important. The stability of the base line would also be affected by displacements attributable to small mechanical movements within the valves*. In this case, too, the maximum usable gain is probably nearer 400,000. This *can* be obtained with 3 stages, and easily with 4.

An additional difficulty with high gain amplifiers of long time constant is 'blocking'. This refers to a lengthy paralysis of the amplifier following the

* Not so much microphony, as movements to relieve stresses caused by the valve warming up.

receipt at the input of a brief pulse of excessive size, and occurs as follows: Suppose the polarity of the pulse is such as to drive the anode of the penultimate amplifier valve positive. The grid of the final valve tries to follow it, but grid current flows to charge the coupling capacitor and increase the P.D. across it. At the end of the pulse the penultimate anode returns to its proper mean potential, carrying the final stage grid far more negative than its proper bias and cutting off the final valve. The amplifier is then paralysed until the P.D. across the coupling capacitor returns to normal, a period which is often about ten times the time constant for the coupling, and may therefore exceed half a minute. In a double-sided circuit a large input pulse of either polarity will block the amplifier. The magnitude of blocking effects depends on the amplifier gain, so we have here another factor to limit this quantity.

The base line produced by direct-coupled amplifiers is subject to flicker and carbon resistor noise, to the effects of small valve electrode movements and to drift. $100 \mu\text{V}$ per in. may be regarded as a reasonable upper limit to the deflection sensitivity in direct-coupled systems.

It should be noted parenthetically that if a record contains a small event, which might be signal and might be noise, a certain time after the delivery to the preparation of a stimulus, and if a similar event appears at a corresponding time upon repeating the stimulus, then the probability that the event is a signal and not noise is increased. This gives us a clue to a technique which allows us to measure signals actually *smaller* than the noise, and which has been exploited by Dawson¹. If corresponding segments of the traces from repeated experiments are added, then the signals sum to produce a total which is directly proportional to the number of experiments, whereas the noise produces a total which rises as the square root. Dawson used a bank of capacitors which were charged, in succession, from the amplifier output via a rotating switch synchronized with the time-base generator. In this manner the voltage across each capacitor corresponds to the average of a large number of values from a particular segment of the record, and a voltage distribution eventually emerges from the whole bank which may be scanned to produce a relatively noise-free record. The degree of freedom from noise depends upon the number of times the experiment has been repeated.

When special techniques such as this can be employed the usable amplifier gain can be increased. The difficulty is, of course, to be sure that each repeated stimulus is producing similar physiological events. If one can be certain that this is so, then theoretically an indefinitely small effect can be detected if one pursues the experiment long enough.

The division of the recording chain into pre- and main-amplifiers is usual and convenient, for these reasons:

(1) *Flexibility*. The nature of the pre-amplifier is determined by what is to be measured. Similarly the type of main amplifier depends on the display device—cathode ray tube, penwriter, etc. More can therefore be done with fewer units.

(2) *Electrical convenience*. Whilst the main amplifier can usually be fed from mains-driven power supplies of conventional type, pre-amplifiers are generally supplied from batteries, or from the mains via special, very highly smoothed and stabilized, power packs.

BIOLOGICAL AMPLIFIERS

(3) *Mechanical convenience.* Most of the controls in frequent use are on the pre-amplifier. The latter can therefore be arranged close to the experimenter, the main amplifier being relegated to some convenient and unobtrusive position. In the Cambridge School of Physiology under Matthews, the experimenter, preparation and pre-amplifier with batteries are all together in a wire-mesh screening cage. The remainder of the electronic apparatus—apart from the RF stimulator probe—is kept outside.

If the display device is to be a cathode ray tube it may be worth while considering buying a commercial oscilloscope, such as the Cossor 1049 or the Nagard DT103. Both these have internal direct-coupled deflection amplifiers, making the provision of a separate main amplifier unnecessary. Similarly, manufacturers of penwriters commonly supply a main amplifier especially designed to go with their products. Pre-amplifiers, on the other hand, have in general to be home made.

The distribution of labour between pre- and main-amplifier is usually arranged so that the signal level is raised to the order of 1 V in the former, further amplification as necessary being carried out by the latter.

A.C. COUPLED PRE-AMPLIFIERS

A very simple single-sided capacitor-coupled pre-amplifier is shown in *Figure 39.1*. This is suitable for recording action potentials from nerves via

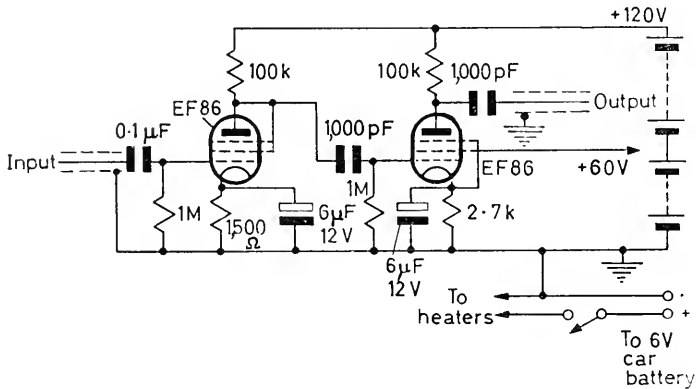


Figure 39.1

low resistance electrodes, and is very suitable for beginners' use. Because it is single-sided, it is unable to discriminate against interfering a.c. fields, and consequently the preparation and electrodes must be totally screened. Unless a screened room is available this rather restricts its application to small animals. As an added measure of protection against 50 cycles pick-up the coupling time constants are chosen so that this frequency lies well below the bottom end of the amplifier pass-band. The first stage is triode connected to eliminate partition noise, and the signal is fed to it via a CR network; this keeps the valve's grid current out of the preparation. The gain lies between 1,500 and 2,000 times; it will depend upon the particular valve specimens in use. The upper limit of frequency response is largely determined by the output cable; it is profitless to have the pass-band excessively

DIRECT-COUPLED PRE-AMPLIFIERS

wide, as response above, say, 10,000 cycles per second, adds little to the signals but contributes unwelcome noise. If we arrange the response to be about 3 dB's down at 10,000 cycles it turns out that the cable capacitance can be 160 pF. This allows one to use 10 ft. of the usual polythene insulated variety, which is quite convenient.

For work on large preparations—including human beings (e.g. cardiography and encephalography)—which are difficult to screen properly, the use of differential amplification becomes essential. A very typical design is given by Dickinson² and is reproduced in *Figure 39.2*. The first stage valves

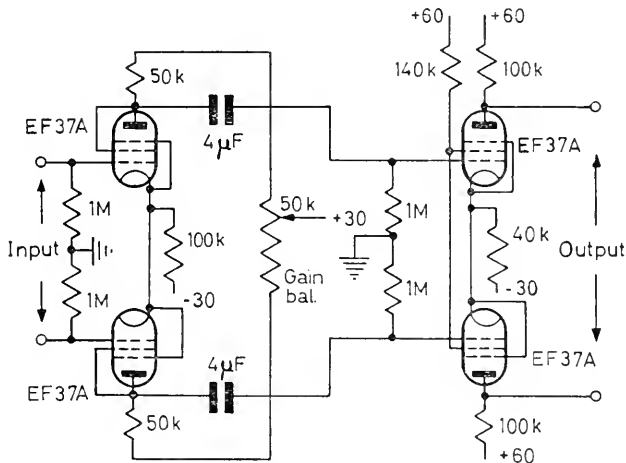


Figure 39.2

are triode-connected and facilities are provided for gain balancing. The time constant of the interstage coupling is 4 seconds, rendering the pre-amplifier suitable for studying slow phenomena such as spinal cord root potentials, electrocardiograms, etc. The heater supply is 6 V, but the first stage heaters are connected in series so that each valve receives only 3. A low cathode temperature and hence low noise are thus obtained.

DIRECT-COUPLED PRE-AMPLIFIERS

As soon as the frequency response of amplifiers has to extend down to zero frequency there is an immediate increase in complication, which seems to be inevitable and makes it worth while to consider carefully before embarking upon it. There are three approaches, each dictated by the requirements.

(1) *'Straightforward' d.c. amplifiers.* These have been discussed in Part I and attention has been drawn to the necessity for distinct static balancing and gain balancing controls and to the existence of drift.

(2) *'Carrier' amplifiers.* These are not true direct-coupled amplifiers but behave as if they were. The input signal is converted into an alternating one ('chopped') at a fixed frequency which is much higher than the highest frequency to which the apparatus is to respond. The alternating voltage thus produced undergoes conventional capacitor-coupled amplification, followed by phase-sensitive rectification. If it is then passed through a low-pass

BIOLOGICAL AMPLIFIERS

filter which cuts off at a frequency below the chopping frequency but above the upper limit of the required pass-band, an amplified copy of the original input is obtained (*Figures 39.3 and 39.4*). Chopper amplifiers are characterized by low drift.

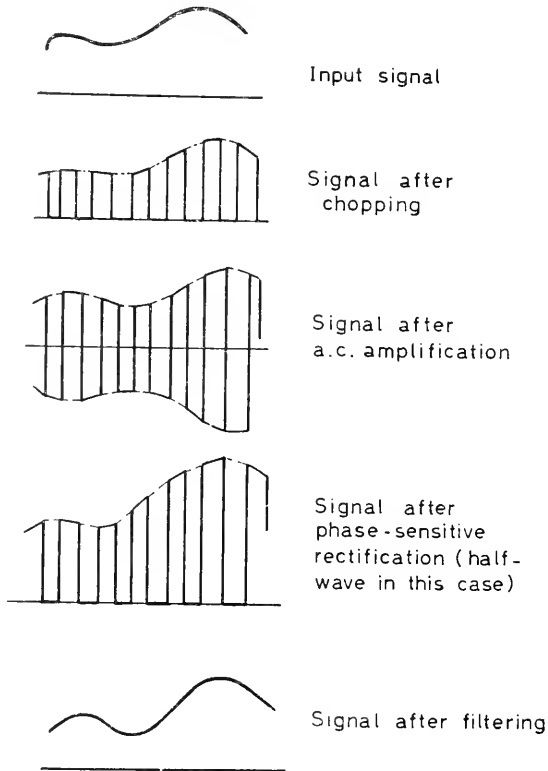


Figure 39.3

(3) *Galvanometer amplifiers.* This technique is due to A. V. Hill. The input signal is applied to a moving-coil reflecting galvanometer and deflects the beam of light across a differential photoelectric cell. The cell output is

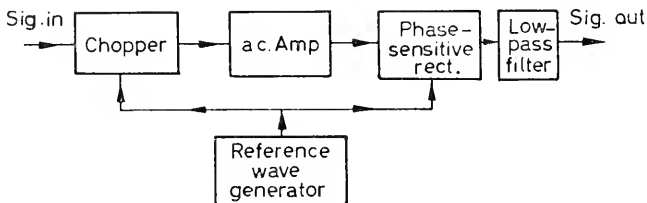


Figure 39.4

amplified by valves and fed back negatively to series oppose the original electrical input (*Figure 39.5*). This has the usual effect of increasing the input resistance of the device—from a hundred ohms or so, the resistance of the moving-coil, to the order of a megohm—and of raising the upper limit of

DIRECT-COUPLED PRE-AMPLIFIERS

the galvanometer frequency response about 10 times. Galvanometer amplifiers are characterized by extremely low noise.

The bulk of electrophysiological work is on exposed nervous tissue, and for this reason the electrical signals which may be drawn off are quite large, roughly in the range $50 \mu\text{V}$ – 50 mV . In these applications the 'straightforward' direct-coupled amplifier is quite satisfactory. Drift presents tiresome but not insuperable difficulties, and in-so-far as the signals are relatively large, the amplifier noise—provided reasonable precautions are taken—is not excessive. In microelectrode work the drift and noise originating in the electrode are in any case far more serious than that arising in the amplifier.

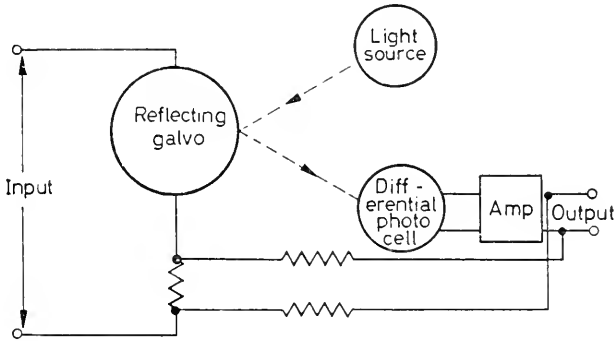


Figure 39.5

There are, however, fields in which much smaller steady potentials have to be measured, usually from low-resistance surface electrodes. A case in point is the study of eye movements by measuring the vertical and horizontal components of electro-oculogram. These are of the order of a few microvolts only. In work on brain or spinal cord activity there is a sense in which an experiment is over in a single traverse of the cathode ray tube trace, perhaps 100 msec; in this time the preparation has been stimulated and the response elicited and recorded. In an eye movement study the time scale is likely to be of the order of minutes. Clearly, then, only a very small drift rate indeed is acceptable. This is the kind of work for which the chopper amplifier is suited.

Galvanometer amplifiers are at present rarities and it seems likely that they will remain so. Their freedom from noise gives them a position of supremacy which is probably unassailable in studies of very feeble electrical effects, but their bulk, necessary for elaborate resilient mounting and sound-proofing, coupled with their use of special photocells and galvanometers of extreme sensitivity, neither of which are readily available, may account for their infrequency of appearance.

Straightforward d.c. amplifiers

An enormous amount of work has been directed towards the design of better d.c. amplifiers, and descriptions of new circuits are published steadily. Unfortunately most of these are developed for purposes other than—and are not suitable for—bioelectric studies. The reader who is interested is referred to a very useful study on the subject by Yarwood and Le Croisette³ which includes a very long bibliography of work in this field.

BIOLOGICAL AMPLIFIERS

Of the d.c. amplifiers which have been published for electrophysiological work, a notable example is due to Bishop and Harris⁴. This might perhaps, without impertinence, be described as the ultimate in biological amplifiers. Both main- and pre-amplifiers are described. The latter comprises a pair of cathode followers, followed by two stages of balanced amplification by double triodes of a type (6J6 or ECC91) in which both valves share the same cathode. The object here is presumably to mitigate the effects of emission fluctuations. These two stages have pentode valves in the common cathode lead to secure a very high rejection ratio. The fourth stage, a pair of pentodes,

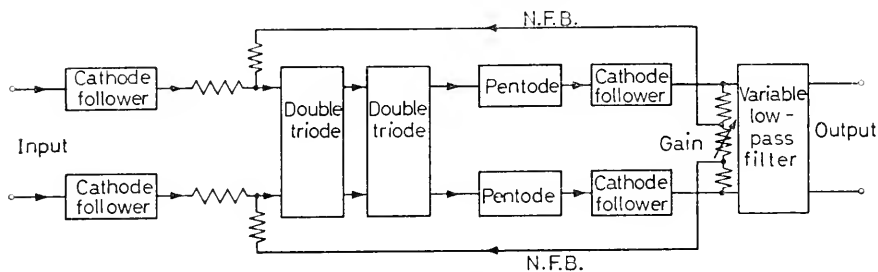


Figure 39.6

provides additional amplification and the fifth is a pair of output cathode followers. From the CF outputs the signal is passed back to the first double triodes to control the gain by variable negative feedback, and is passed forward to the main amplifier via a single stage of low-pass filtering in which the capacitors are switched to limit the upper frequency response as required (Figure 39.6). The pre-amplifier includes a stabilized HT supply and uses 15 valves in all (including 1 stabilizer tube).

Not all physiologists would agree that so much complexity is desirable, though there is no question that this amplifier is capable of a very fine performance. A much simpler amplifier has been described by Asher⁵ which possesses the valuable feature that it may be switched to capacitor-coupled operation as required. An unusual feature is the absence of any division into pre- and main sections; a pair of cathode followers is followed by 4 stages of balanced pentode amplification, after which the signal is of sufficient magnitude to deflect a cathode ray beam. The whole amplifier is supplied from batteries. It is of the 'climbing' type and 240 V HT are required. A mains-driven unit in which beam-tetrodes arranged as cathode followers drive a penwriter is also described.

A pre-amplifier designed by the author for general purposes, including microelectrode work, is shown in Figure 39.7. Before discussing it we will consider for a moment the question of input cathode follower grid current in direct-coupled apparatus.

Consider what happens when a microelectrode pierces a hypothetical spherical cell of radius 10μ . The surface area A of the cell is $4\pi r^2 = 12.6 \times 10^{-6} \text{ cm}^2$.

Taking the resistance of cell membrane, ρ , as $1,000 \Omega \text{ cm}^2$, we see that the resistance of the membrane of this particular cell is

$$\frac{\rho}{A} = \frac{10^3}{12.6 \times 10^{-6}} \approx 8 \times 10^7 \Omega$$

DIRECT-COUPLED PRE-AMPLIFIERS

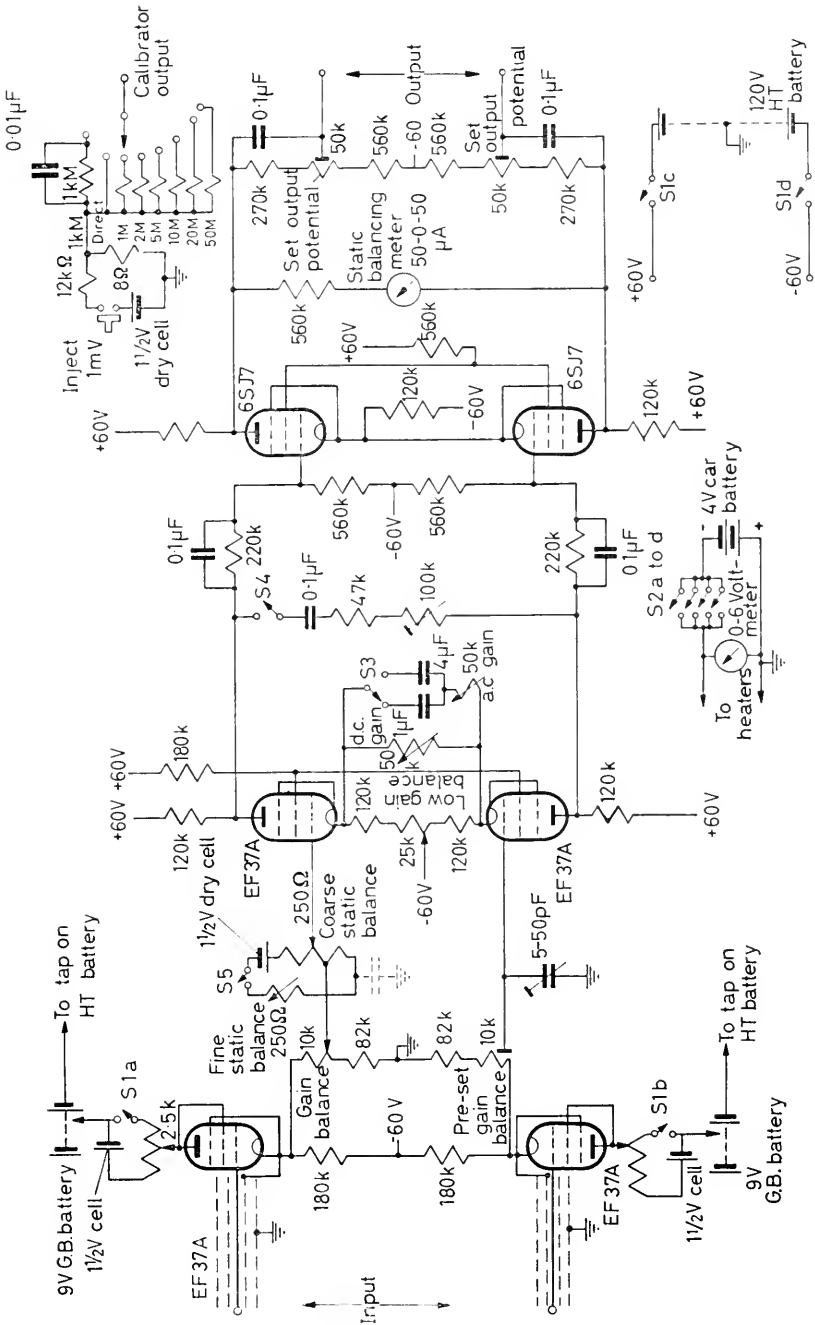


Figure 39.7

BIOLOGICAL AMPLIFIERS

If no special attention is paid to the input cathode-follower grid current, this figure may well be 10^{-9} amps. The effect of introducing a current of this magnitude into the hypothetical cell is to produce a spurious potential difference across the membrane of $(8 \times 10^7) \times 10^{-9}$ V, or 80 mV. Clearly the circumstances of recording are not very physiological.

One often hears or reads statements such as 'the cathode follower valve was selected for low grid current'. Reference to *Figure 39.8*—which shows the form of the grid-current characteristic for a thermionic valve—makes it clear that any valve may be made to pass zero grid current by choice of

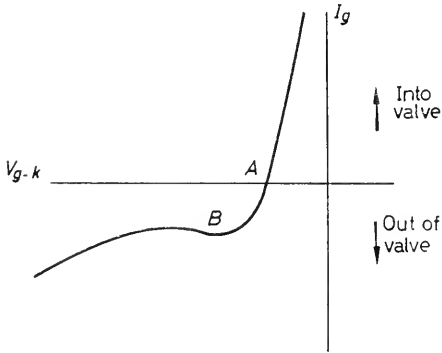


Figure 39.8

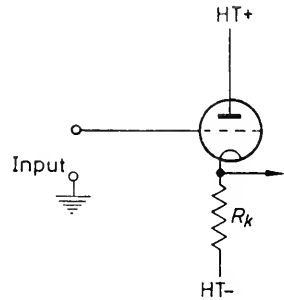


Figure 39.9

appropriate bias (point *A*), but that at this point the input resistance, equal $(\delta V_{gk})/(\delta I_g)$ is rather low. A more usual working point in capacitor-coupled circuits, where the grid current is diverted down the grid leak, would be point *B*, where the input resistance of the valve is high.

Experience shows that if a suitable cathode follower valve is operated under 'low-noise' conditions, reduced heater voltage and low anode current, the order of input resistance corresponding to point *A* working is nevertheless sufficiently high to be large compared with a microelectrode of resistance 50 megohms or so. It appears that for this particular application point *A* rather than the more usual point *B* working is satisfactory.

Consider the direct-coupled triode cathode follower circuit in *Figure 39.9*. The current through the valve is determined by the ratio of the negative supply voltage to R_k , since the cathode and grid are at approximately earth potential. The grid-cathode bias depends on the positive anode supply, since the anode-cathode and grid-cathode voltages are related by the rather constant parameter μ . It emerges, then, that the grid current can be set to zero by appropriate choice of HT+.

This makes it a good deal easier to choose a suitable specimen for this important position. Instead of rejecting valves out of hand because their zero grid current point does not occur at the particular value of HT+ (arbitrarily chosen) which is in use, one can concentrate on selecting for low noise and microphonicity. If a number are found which are satisfactory in these respects, efforts can be directed towards choosing one whose grid current characteristic cuts the $I_g = 0$ axis at the minimum slope; this specimen will then hold its low grid current setting best.

DIRECT-COUPLED PRE-AMPLIFIERS

In the case of pentode cathode-followers the bias is determined by the screen potential, and it is this which must be varied to find the zero grid current condition.

The range of anode voltages which it is necessary to provide for triode cathode followers, and of screen voltages for pentode versions, lies between

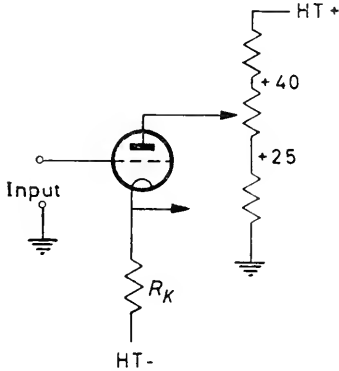


Figure 39.10

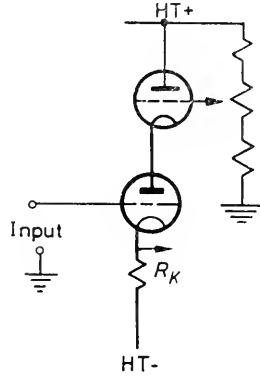


Figure 39.11

+25 and +40 V for the types of valve tested—Mullard EF37A and ME1400. The author has not tried this technique with pentodes, as a variable supply carried on the cathode is not easy to arrange. With triodes, however, it is not too difficult; the first method which occurs to one is *Figure 39.10*, but

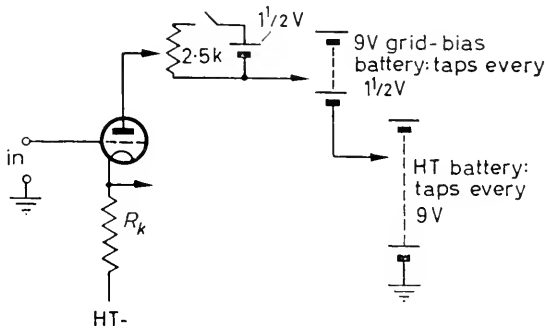


Figure 39.12

this is not satisfactory. For correct operation, the resistance seen looking back into the network supplying HT to a cathode follower ought to be low; to achieve this the resistor values in the divider chain have to be so small that a current is wasted down them of a magnitude which is intolerable in battery-operated gear. A promising alternative is to feed the cathode follower from another cathode follower (*Figure 39.11*). In this case the output impedance of the variable HT supply to the lower valve is $1/g_m$ for the upper, and the potential-divider chain can have as high a value as is wished. However, in view of the extra valve involved, the method finally adopted is shown in *Figure 39.12*. The required HT is selected to the nearest 9 V by

BIOLOGICAL AMPLIFIERS

taps on the HT battery; then to the nearest $1\frac{1}{2}$ V by taps on the grid bias battery. Final adjustment to within a few millivolts is performed on the 2.5 k ohms potentiometer, which is supplied from a large unit cell (U2). In this manner the output impedance of the supply is never greater than about 1.25 k ohms if the batteries are in good condition. The price paid for the use of a triode cathode follower, rather than pentode, is an input capacitance of about 4 pF instead of some $1\frac{1}{2}$ pF.

In the complete amplifier (*Figure 39.7*) the interstage and output coupling networks are seen to be of the potential-divider type and about $\frac{1}{3}$ of the signal at zero and low frequencies is lost in each. However, the upper resistors of the potential dividers are shunted by capacitors so that fast alternating components are passed without attenuation. Neglecting for the moment the network associated with the first amplifier stage cathodes, the frequency

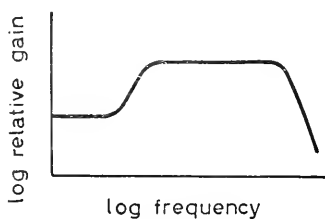


Figure 39.13

response of the amplifier is clearly of the form of *Figure 39.13*; the full gain is available for the detection of action potentials which may be distant from the microelectrode, and rather less than half for measuring membrane potentials and slow changes such as root potentials. In practice this is usually precisely what one wants, since the latter two are large, and in addition the reduced gain at zero frequency gives less trouble from drift. However, for workers who find such a frequency response confusing, the 'high frequency lift' may be approximately cancelled by introducing a series CR circuit with the switch S4; this is simpler than switching out the 4 capacitors in the coupling circuits.

The arrangement of the 'a.c. gain' and 'd.c. gain' controls is due to Matthews⁶. These terms are convenient but not quite accurate, since the d.c. gain control actually operates at all frequencies, therefore the d.c. gain cannot be more than the a.c. gain; in practice this presents no difficulty because one never wants it to be. The network enables one to reduce d.c. gain steadily until the amplifier behaves for practical purposes as if it were a.c. coupled. This is a much more flexible arrangement than one which is switched between 'pure a.c. coupled' and 'pure direct coupled'.

Meters are provided to facilitate balancing of the amplifier and to check the state of the heater battery. In addition a simple calibrating device enables one to:

- (a) check the gain (600 max on d.c., 1,300 max on a.c., single-sided output)
(1,200 max on d.c., 2,600 max on a.c. double-sided output)

- (b) estimate the microelectrode resistance or resistances;

- (c) set the input cathode follower grid current to a low value and measure it.

The output terminals are fed in push-pull and both may be used to feed a

DIRECT-COUPLED PRE-AMPLIFIERS

main amplifier which possesses a differential input. For those which are not, i.e. most commercial oscilloscopes, a single-sided output can be taken from either of them. There will be some loss in rejection ratio, and the gain will of course be halved.

Although this particular pre-amplifier design employs obsolescent valve types, the setting-up procedure will be described because it is quite typical and serves to indicate the line of approach for differential direct-coupled amplifiers in general. The procedure is as follows:

(1) Earth both cathode follower inputs and turn the d.c. gain control to maximum. Check on the output meter that the amplifier can be balanced by manipulation of the coarse and fine balance controls. Keep it balanced in this manner during operations 2-5.

(2) Disconnect one cathode follower input from earth and connect it instead to the calibrator output terminal. Connect a measuring oscilloscope to the amplifier output. Inject a 1 mV pulse and measure the d.c. gain. Open S_4 , set 'a.c. gain' to maximum, replace the calibrator by a sine wave oscillator set to 1,000 c/s, 1 mV R.M.S., and from the oscilloscope check the a.c. gain.

(3) Switch the oscillator output to square waves of about 1 mV amplitude, 200 c/s, close S_4 and adjust the 100 k pre-set potentiometer until the reproduced waveform is as near as possible square. Open S_4 again.

(4) Connect both cathode follower inputs together and inject about 100 mV R.M.S. at 50 c/s between them and earth. Turn the gain balance control until the minimum output appears on the oscilloscope. This gets the rejection ratio as high as possible.

(5) Increase the frequency to 5,000 c/s and turn the trimming capacitor until the minimum output appears on the oscilloscope. This compensates for the capacitance to earth of the static balancing circuit.

(6) Short the inputs to earth again and carefully balance the amplifier. Turn the d.c. gain control to minimum and re-balance using the 'low gain balance'. Return the d.c. gain to maximum and re-balance using the fine and coarse balance controls. The amplifier should now remain in balance irrespective of the d.c. gain control setting.

(7) With the aid of a high resistance voltmeter adjust the output potential controls so that both outputs are at the same potential as earth when the amplifier is balanced; i.e. make sure that the outputs are not superposed upon a steady bias.

(8) Grid current adjustment. Unshort the input to the cathode follower to be set up and connect it to the calibrator output. Switch calibrator to 'direct' and balance the amplifier. Now switch calibrator to 'kM Ω '. The grid current will flow along the kM Ω resistor, developing across it a potential difference which will unbalance the amplifier. Return the calibrator switch to 'direct' and alter the appropriate cathode follower HT supply in such a direction as to off-balance the amplifier in the opposite sense to that in which it was moved by the grid current. Re-balance using the fine and coarse balance controls. Repeat the procedure until inserting the kM Ω into the cathode follower grid circuit produces little or no unbalance. The grid current is given by the voltage at the input corresponding to the unbalance produced, divided by $10^9 \Omega$, and should be less than 10^{-12} amps.

BIOLOGICAL AMPLIFIERS

(9) Microelectrode resistance estimation. Connect the microelectrode to the cathode follower input and dip it into a vessel of earthed Ringer's fluid. Apply 1 millivolt to the cathode follower input terminal from the calibrator (switched to 'direct') and observe the deflection produced on the cathode ray trace. Now operate the calibrator switch to insert resistance in series with the calibrator output until the deflection produced on the trace is halved. The amount of series resistance inserted is then equal to the microelectrode resistance (Figure 39.14).

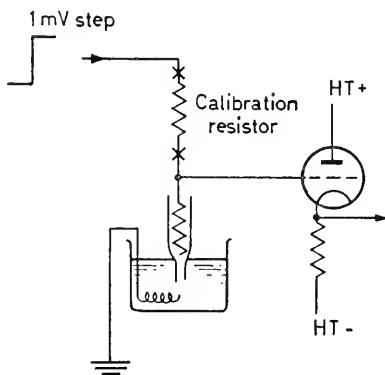


Figure 39.14

Other calibrators—The calibrating device just described is not very accurate, depending as it does on the freshness of the cell. Moreover it is useful to have other calibrating potentials beside 1 mV available. A more sophisticated calibrator is shown in Figure 39.15. This gives up to 1 V in 1 mV steps, but the scheme may be extended to give smaller increments. The pre-set potentiometer is used to set the circulating current to exactly 1 mA as indicated on the meter. The 10 Ω potentiometer is of course continuously variable but has a scale marked 0–10. The assumption is that operating this potentiometer does not affect the circulating current. Clearly with a total circuit resistance of about 1,500 Ω this is substantially true; the error is less than 1 per cent and the meter cannot be read to better than this. Purists may open the circuit at a point such as 'X' and insert another 10 Ω variable resistor, ganged to the first and arranged so that as the first increases the new one decreases.

An even more accurate calibrator is used by the Rockefeller unit of the Physiological Laboratory in Cambridge in which the output can be checked against a Weston's standard cell (Figure 39.16). Since the e.m.f. of a Weston cell is 1.0186 V, if the circulating current be adjusted until there is no microammeter deflection on pressing the check button, then it must be equal to 100 μ A.

Carrier amplifiers

We have seen that these comprise a d.c.–a.c. converter or *modulator*, an a.c. coupled amplifier and a phase sensitive rectifier. The modulator may be mechanical, in which case it comprises a vibrating relay or capacitor, or it

DIRECT-COUPLED PRE-AMPLIFIERS

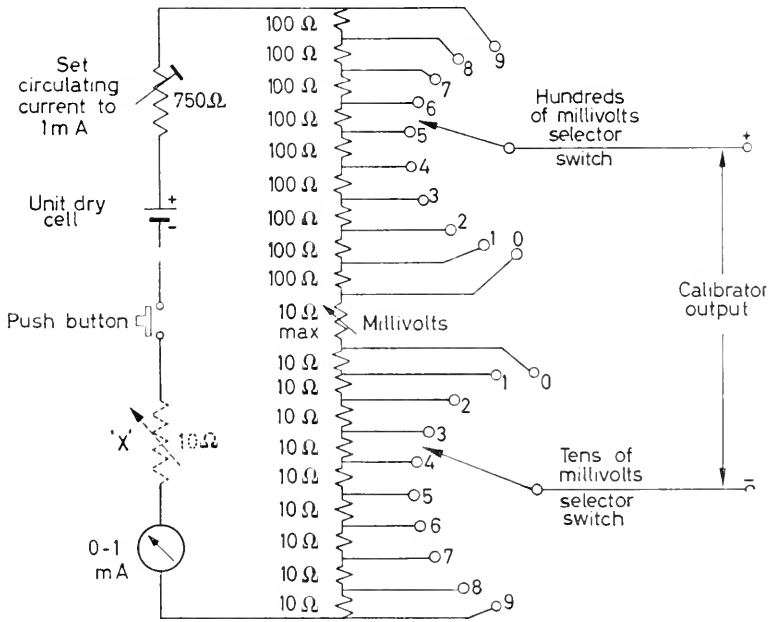


Figure 39.15

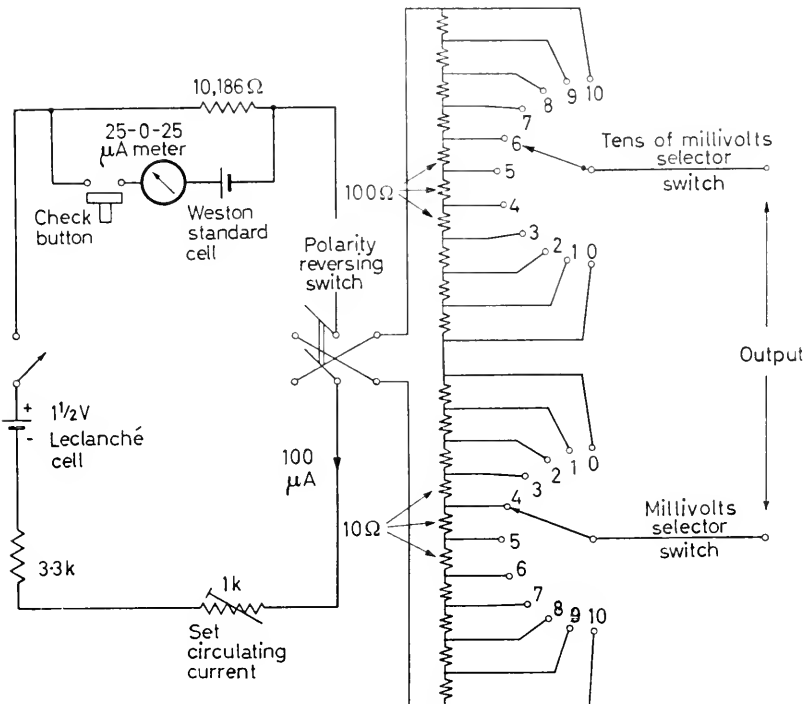


Figure 39.16

BIOLOGICAL AMPLIFIERS

may be purely electronic. Electronic modulators have not been dealt with in Part I because they have little relevance to the subject of this book, but the reader who is interested will find them discussed in any textbook on radio.

Vibrating relay or 'Chopper' amplifiers—Provided the gain required is not too high a simple form of chopper amplifier may be made using an existing a.c. coupled amplifier, a high-speed change-over relay and an oscillator of sufficient power to operate it. A possible scheme is shown in *Figure 39.17*,

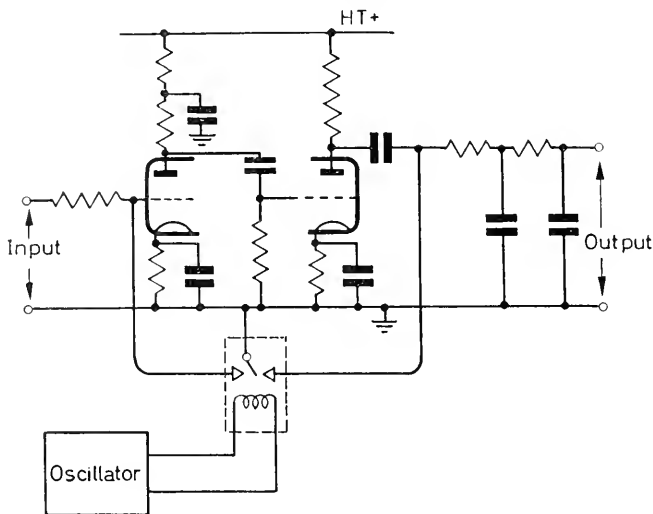


Figure 39.17

where the amplifier is a simple two-stage affair using a double triode. The relay moving contact 'buzzes' back and forth between the fixed contacts, alternately earthing the input and the output. The periodic short-circuiting of the input converts it into a square wave of amplitude proportional to

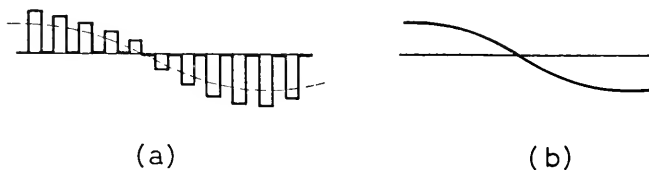


Figure 39.18

itself and of phase depending on its polarity. This emerges from the amplifier about a thousand times greater and is phase-sensitive-rectified by the other relay contact in a manner similar to the Cowan bridge described in Part I. The output is now of the form shown in *Figure 39.18a*, and if this is filtered it has the appearance of *Figure 39.18b* and is a magnified copy of the input. The chopping frequency should be the highest at which the relay will work properly, and the filter can then be designed to cut off at, say, one-third of this, and the amplifier will then handle input frequencies up to, say again, one-tenth.

A difficulty with chopper amplification of this rudimentary type is this:

DIRECT-COUPLED PRE-AMPLIFIERS

Having got a relay to chop properly—no mean feat*—it is tempting to get *all* the necessary gain by a.c. amplification and subsequent phase-sensitive rectification, i.e., to eliminate the subsequent ‘straightforward’ main amplifier. In practice this is usually done, but if the input and output terminals of a high gain amplifier are brought to within a few mm of one another at the relay contacts, the consequent capacitive feedback will almost certainly cause the amplifier to oscillate.

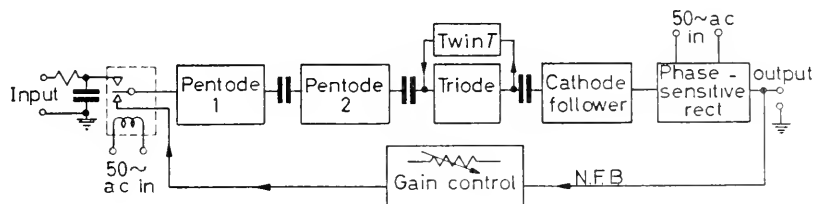


Figure 39.19

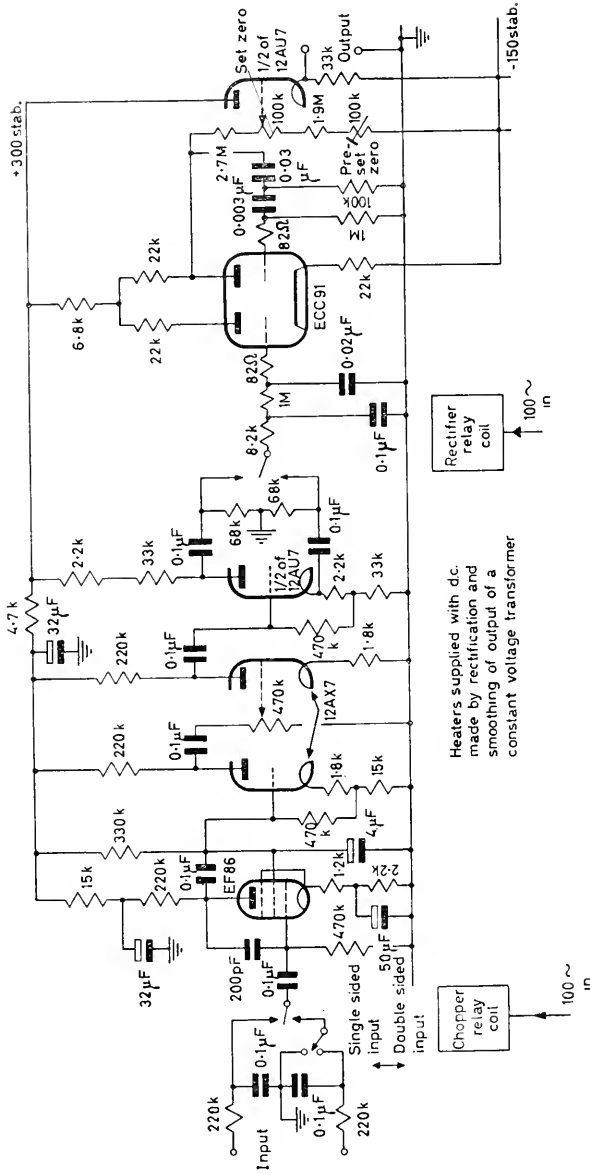
There are two possible ways out of this. One is to use two relays, widely separated, whose coils are fed from the same oscillator, one to short-circuit the input and one to short-circuit the output. The second is to use a different kind of phase-sensitive rectifier altogether.

An amplifier using the second expedient has been described by Nielsen and Rosenberg⁷. Though developed for pH measurement there seems to be no reason why this should not be used for biological studies of slow phenomena; the chopper frequency is 50 c/s, so a frequency response of more than a cycle or two per second would not be expected. The short-term drift is 10 μ V and the long-term (12 hours) 30 μ V. The maximum gain is not stated but is probably well over 100,000. An extremely elaborate electronically stabilized power supply for both HT and heaters is also described. *Figure 39.19* shows the amplifier proper in schematic form. The chopper relay is succeeded by two stages of pentode amplification, after which there is a triode arranged as a 50 c/s tuned amplifier by a negative feedback loop containing a twin *T* network. This restricts the amplifier band-width and hence excludes from the output much of the noise generated at other frequencies by the earlier stages. The phase-sensitive rectifier is of the full-wave type comprising a ring of 4 diodes. Control of gain is by negative feedback to the spare relay contact; as usual the use of negative feedback in this way serves also to increase the input resistance of the device, the expression for which is rather complicated. The low-pass filter in the input circuit could be dispensed with for many applications.

A two-relay chopper amplifier is shown in *Figure 39.20*. This apparatus was developed for electro-oculography by E.M.I. Electronics Ltd., to whom I am grateful for allowing me to reproduce the circuit. The amplifier accepts either single-sided or differential input. After chopping at 100 c/s, there are 3 stages of amplification, followed by a concertina phase-splitter which, with the second relay, forms a full-wave phase-sensitive rectifier. Next comes two stages of low-pass tapered RC filtering, then a fourth filter of the phase-shift feedback type, and lastly a cathode-follower output stage. The gain of

* There must be no ‘contact bounce’.

BIOLOGICAL AMPLIFIERS



Heaters supplied with d.c. made by rectification and smoothing of output of a constant voltage transformer

Figure 39.20

DIRECT-COUPLED PRE-AMPLIFIERS

the amplifier is not less than 10,000 and the drift is equivalent to a signal of less than $\pm 20 \mu\text{V}$ in half an hour. With the input terminal shorted with a $10 \text{ k}\Omega$ resistor, the noise is less than $10 \mu\text{V}$ peak-to-peak [*sic*]. The input resistance is $1 \text{ M}\Omega$. The high-frequency performance is stated in terms of the rise time for a step function at the input, and is given as 100 milliseconds.

The vibrating capacitor amplifier—The possible upper limit of frequency response of chopper amplifiers is set by the upper frequency at which it is possible to get relays to chop satisfactory. At present chopper relay designers seem to be aiming at 50 or 60 cycles operation, as the mains can then be used

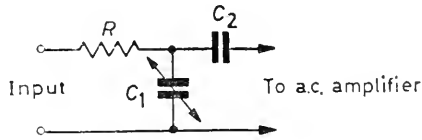


Figure 39.21

to drive them. The vibrating capacitor amplifier promises much faster operation because the mechanical part is merely a metal reed oscillating near a fixed plate. The principle of the method is shown in Figure 39.21. C_1 is the vibrating capacitor and C_2 is for coupling. R is a high resistance such that RC_1 greatly exceeds the period of one vibration. Then if the input is a relatively slow-changing voltage, the charge on C_1 is also slow-changing and proportional to it, and the voltage across C_1 is proportional to the reciprocal of the capacitance and to the charge. Therefore the voltage across C_1 contains an alternating component at the vibration frequency proportional to the input.

The performance of amplifiers employing mechanical modulators has been reviewed by Chance⁸ who concludes that chopper amplifiers are more satisfactory with low impedance inputs and vibrating capacitor types with high. It is thus not surprising that designs which have appeared for the latter have been rather for the measurement of extremely small currents from, e.g. ionization gauges^{9,10}. Nevertheless it is to be hoped that before long the vibrating capacitor amplifier will be tried for biological amplification. The amplifier of Thomas and Finch is particularly interesting as the vibration frequency is 550 c/s; there is also a novel type of phase-sensitive rectifier, using a double triode valve.

Amplifiers with electronic modulation—The great advantage here is that the carrier frequency can be sufficiently high for the signal pass-band to be as wide as one could wish. Thus an amplifier described by Lampitt¹¹, by employing a carrier of 50 kc/s, could handle signals from zero to 5,000 c/s. The gain was 100,000 but the equipment was designed to measure signals of the order of one mV, so there is no guarantee that noise level was not rather high. The rectification of Lampitt's amplifier was not phase sensitive, so that the output could only be of one polarity, which is somewhat disadvantageous. In addition, no component values were given, so that the production of a similar amplifier would require development almost from scratch.

Galvanometer amplifiers—For applications requiring superlative performance at low frequencies and moderate input impedance, the reader is referred to Preston¹², Hill¹³, and Frankenhaeuser and Macdonald¹⁴.

BIOLOGICAL AMPLIFIERS

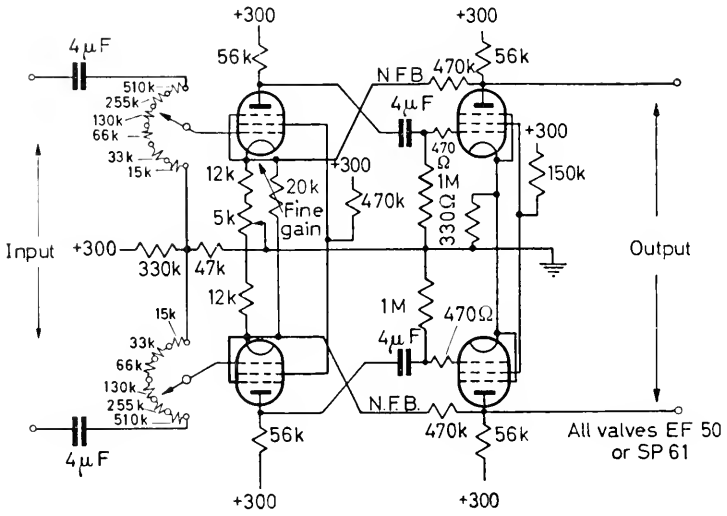


Figure 39.22

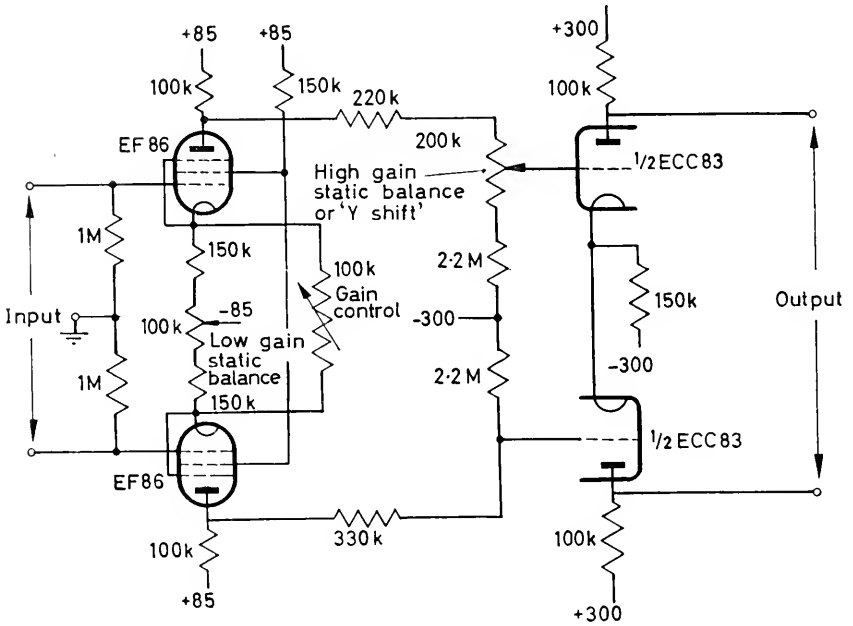


Figure 39.23

MAIN AMPLIFIERS

Main amplifiers for electrostatic cathode ray tube deflection

A capacitor-coupled main amplifier having a time constant of 2 seconds which is in every way typical is given by Dickinson¹⁵ and is reproduced in *Figure 39.22*. The amplifier accepts a balanced input and delivers a balanced output to the deflector plates. Coarse (stepped potential divider) and fine (variable negative feedback) gain controls are provided, and Dickinson points out that it is vital that the resistors used in the coarse control divider chain be accurately matched lest an in-phase signal applied to the input develop an out-of-phase component and appear on the tube base line as an artefact. A direct-coupled counterpart due to the author is shown in *Figure 39.23*. Despite its simplicity the gain is such that with a 6 in. cathode ray tube at an EHT of 2 kV, the spot is deflected from top to bottom of the screen by 100 mV. The frequency response could be improved by adding compensating capacitors to the interstage potential dividers, but the performance is satisfactory for examining the shape of an amphibian action potential. It must be admitted that the form of *Y* shift control used here acts also as a differential gain control and can theoretically spoil the anti-in-phase rejection ratio of the recording chain. In practice, the effect is not serious, since the control follows 3 (if the pre-amplifier is included) stages of accurate differential amplification. By the time signals reach it, the effect of in-phase components has been substantially suppressed.

Main amplifiers for magnetic cathode ray tube deflection

In order to demonstrate electrophysiological phenomena to students an oscilloscope possessing a large screen is a valuable asset. Large-screen electrostatic cathode ray tubes can be made to special order but naturally they are expensive and in addition they are rather long, leading to an instrument of unsatisfactory dimensions. Television tubes are quite satisfactory for low-frequency (0–10,000 c/s) oscilloscopy, are cheap and, because a wide deflection angle is possible with the magnetic system, physically short.

Magnetic deflection circuits are quite straightforward, but consume rather a lot of current. It is necessary that the back e.m.f. occurring across the deflector coils should never be so great as to carry the anodes of the output valves outside their working region; thus the coil self-inductance must be restricted. The self inductance depends on the square of the number of deflector coil turns, whereas the deflecting capability of the coil is proportional to the number of turns and, of course, the deflector coil current. Therefore to scan completely the tube face at a high frequency requires coils with a rather small number of turns, and a large current. Magnetic cathode ray tubes and deflector coils have been discussed in Part III, and the circuit of a simple practical amplifier devised by the writer given (*Figure 32.39*). The output valves operate in class AB push-pull. The standing anode current for the four is adjusted to 60 mA total in the absence of any input, and the trace centred. Upon the arrival of a direct input signal the current in one or other pair rises to a maximum of 200 mA, at which point the spot is at the top or bottom edge of the screen. Damage through overloading is

BIOLOGICAL AMPLIFIERS

impossible, because if the input rises further, the output stage screen current becomes so high that the stabilizer tube goes out. When this happens there is heavy negative screen feedback and the anode current ceases to rise any further.

Loudspeaker amplifier

Many workers like to be able to hear the result of an electrophysiological experiment as well as see it on an oscilloscope; when the eyes are engrossed

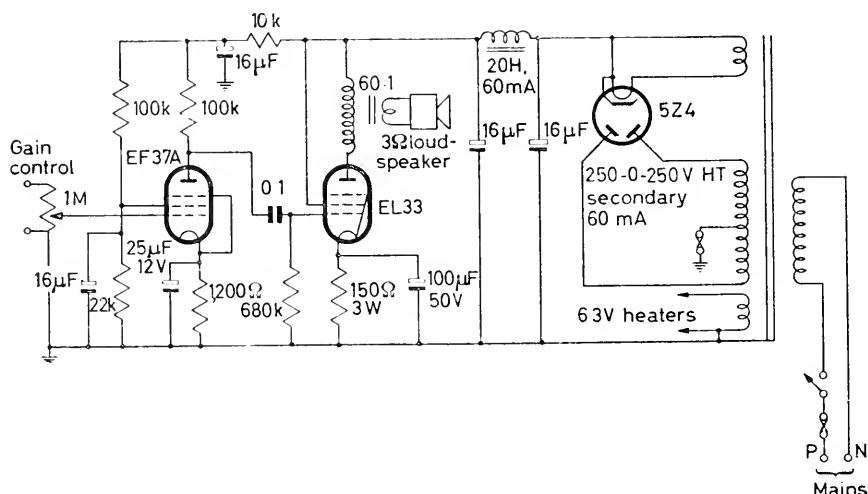


Figure 39.24

in peering down a dissecting microscope at the site of the recording electrode tip, auditory information about what is being recorded may be extremely welcome. A straightforward audio-frequency amplifier and loudspeaker enable one to hear action potentials, and in addition there is a sudden change in noise level at the instant when the microelectrode tip, which is usually invisible, first makes contact with the tissue. This reduction signals the point at which to zero the micromanipulator dial gauges so that henceforward they are direct reading in penetration.

A simple loudspeaker amplifier and power pack are shown in *Figure 39.24*. This is another unit recommended for beginners, and in conjunction with the single-sided pre-amplifier, *Figure 39.1*, enables a start to be made in electrophysiological work.

The bleater

Useful as the above amplifier is, it cannot signal the magnitude or polarity of a steady potential at a microelectrode tip; on piercing a muscle or nerve cell membrane the loudspeaker would merely emit a 'plop'. A device which does, which is not really an amplifier at all but which it seemed appropriate to mention here, is the 'bleater', first reported by Draper and Weidmann¹⁶. A possible circuit is shown in *Figure 39.25*. A thyatron relaxation oscillator has its frequency controlled by a bias derived in part from a battery and

REFERENCES

part from the upper cathode ray tube Y deflector plate. A small loud-speaker is included in the capacitor discharge path. A rather lugubrious musical note is emitted at a pitch which depends on the level of the cathode ray tube trace. With practice it is possible to make a good estimate of the latter by listening to the sound.

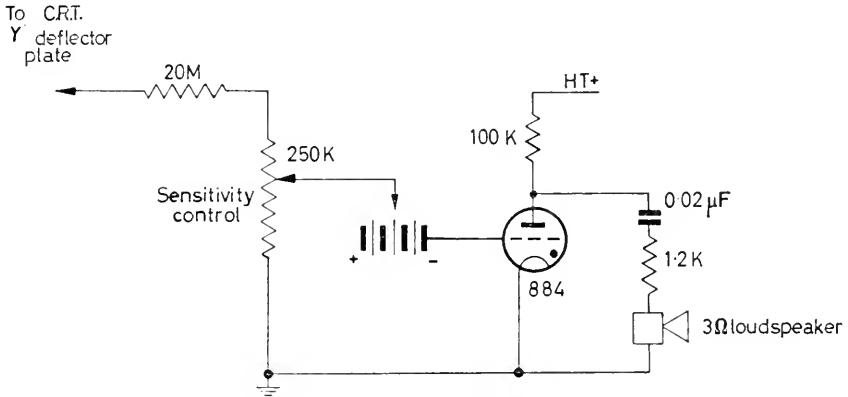


Figure 39.25

Penwriter amplifiers

Penwriters are usually bought, and the manufacturers commonly supply a main amplifier specially designed to go with them. The transient and frequency response of penwriters depends upon the damping placed across them by the output resistance of the amplifier; therefore the two have to be designed together. Readers who are interested in home-made types may care to consult Grey Walter and Brooks¹⁷ (who describe an entire E.E.G. apparatus) and Tanner and Harrington¹⁸.

REFERENCES

- ¹ DAWSON, G. D. *Electroenceph. Clin. Neurophysiol.* 6 (1954) 65
- ² DICKINSON, C. J. *Electrophysiological Technique*. Electronic Engineering
- ³ YARWOOD, J. and LE CROISSETTE, D. H. *Electron. Engng.* 26 (1954) 14
- ⁴ BISHOP, P. O. and HARRIS, E. J. *Rev. sci. Instrum.* 21 (1950) 366
- ⁵ ASHER, H. *Electron. Engng.* 23 (1951) 170
- ⁶ MATTHEWS, B. H. C. *J. Physiol.* 122 (1953) 2P
- ⁷ NIELSEN, S. O. and ROSENBERG, T. *J. sci. Instrum.* 31 (1954) 401
- ⁸ *Waveforms*. Radiation Laboratory series, McGraw-Hill
- ⁹ PALEVSKY, H., SWANK, R. K. and GRENCHIK, R. *Rev. sci. Instrum.* 18 (1947) 298
- ¹⁰ THOMAS, D. G. A. and FINCH, H. W. *Electron. Engng.* 22 (1950) 395
- ¹¹ LAMPITT, R. A. *Electron. Engng.* 18 (1946) 347
- ¹² PRESTON, J. S. *J. sci. Instrum.* 23 (1946) 113
- ¹³ HILL, A. V. *J. sci. Instrum.* 25 (1948) 225
- ¹⁴ FRANKENHAEUSER, B. and MACDONALD, D. K. C. *J. sci. Instrum.* 26 (1949) 145
- ¹⁵ DICKINSON, C. J. *Electrophysiological Technique*. Electronic Engineering
- ¹⁶ DRAPER, M. H. and WEIDMANN, S. *J. Physiol.* 115 (1951) 74
- ¹⁷ GREY WALTER, W. and BROOKS, A. A. *Electron. Engng.* 19 (1947) 221
- ¹⁸ TANNER, J. A. and HARRINGTON, B. G. V. *J. sci. Instrum.* 28 (1951) 33

METHODS OF RECORDING

The traditional physiological recording device is, of course, the smoked-drum kymograph. The writing points may be operated mechanically by cords, pneumatically by tambours, hydrostatically by floats, or electrically by a variety of electromagnetic devices. Advantages of the kymograph are that it is simple, the recording medium is cheap, and the results of extremely long experiments can be displayed in records of manageable size. By the same token, however, it is inappropriate for use in the field for which electronic techniques are so admirably suited—the display of rapid phenomena. In general the recording of quantities electronically derived is achieved by other means, though Keynes¹ has employed a kymograph to register counts from radioactive isotopes via an electronic scaler.

Within the frequency range 0–100 c/s the electromagnetic penwriter is satisfactory. Penwriters may be of the moving-coil or moving-iron variety, and the record may be made in a number of ways. In the most usual, ink is fed under hydrostatic pressure along a thin tube to a nib at the end of the writing arm. In this case the writing medium is ordinary paper and is therefore cheap, but difficulty may be experienced from clogging of the supply pipes, and an accidental overdriving of the recorder can cause widespread distribution of ink. In another method the writing point is a hot stylus which discolours a special paper as it passes over it. A third is entirely electrical and uses ‘Teledeltos’ paper; a current of about 50 mA is passed from the writing point into the paper and away via a metal table which supports it. At the point of entry of the current the paper is burned, leaving as a trace a fine black line. The latter two methods are extremely reliable, but the special papers required are naturally expensive.

The upper limit of frequency response of penwriters is set by the moment of inertia of the mechanical system, particularly of the writing arm, which has to be long if the machine is to write in anything like rectangular coordinates. In one ingenious commercial recorder the writing arm is eliminated and is replaced by a jet of ink which is sprayed from a nozzle which rotates with the mechanical system. In this manner the upper frequency limit has been raised to 500 c/s.

Another commercial recorder has a frequency range which—for biological work—is for practical purposes unrestricted, but this is obtained at the cost of considerable uncertainty in the accuracy of any particular fragment of the record. The input signal is used to control the frequency of a high-frequency oscillator, whose output feeds a battery of band-pass filters to each of which is connected one of a battery of fixed styli, which mark the paper when excited (*Figure 40.1*). Thus any particular instantaneous signal voltage causes oscillations of a definite frequency, and these are passed by one of the filters and cause the appropriate stylus to mark the paper. By choosing a sufficiently high mean oscillator frequency, the

METHODS OF RECORDING

signal frequency may be made as high as required. The disadvantage of the system lies in the finite number of styli and hence number of possible 'quantized' values the record may have.

The bulk of electrobiological recording above 100 cycles, and much that occurs at lower frequencies, is performed by photographing the face of

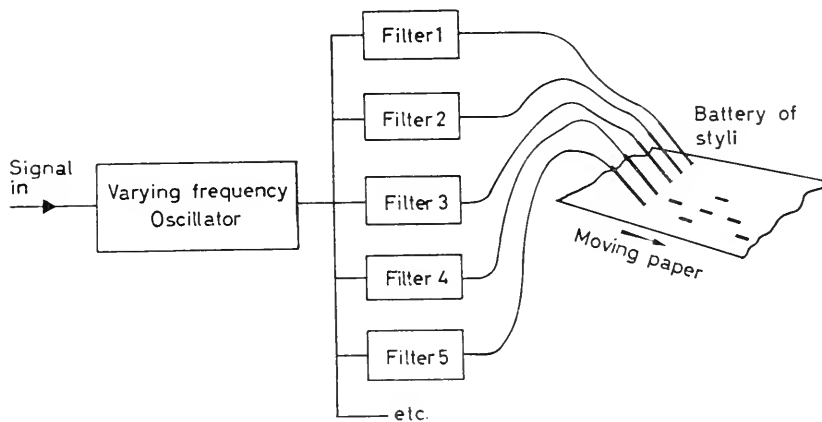


Figure 40.1

a cathode ray tube. The method is accurate, the record can be read with precision, and the frequency range is practically unlimited. A disadvantage is that the record is not at once available for inspection, though a machine has been described which presents a developed photographic record only one second after exposure^{2,3}. 35 mm film or paper is usual: the former is generally perforated for sprocket engagement, which restricts somewhat the useful recording area but ensures a positive drive, while paper is often unperforated and relies for pulling through the camera on friction rollers, which sometimes slip. Two ways of photographing the trace are possible, which might be termed the electrical time-base, and the mechanical time-base, methods.

The electrical time-base method—This is suitable for recording phenomena which are completed within, say, 500 msec of a definite stimulus. The stimulus is delivered regularly and triggers an electronic time-base generator such that the spot is swept across the cathode ray tube face, the response being applied in a direction at right angles in the usual manner. Film can be drawn through the camera at a constant speed in a direction also at right angles to the time-base, so that the finished record might have the appearance of Figure 40.2.

A difficulty arises when the sweep time becomes comparable with the interval between stimuli, or when a 'continuous-running' time-base is used, for the record then assumes the somewhat unsatisfactory appearance of Figure 40.3. This is because, on raising the stimulus rate, the film speed has to be increased to prevent crowding of successive sweeps on the record. When this is done the film speed becomes comparable with the cathode ray tube spot sweep speed, producing the type of distortion shown. It is possible

METHODS OF RECORDING

to correct for this by injecting a suitable fraction of the time-base generator output into the deflection amplifier, but it must be remembered that a different fraction is required at each change of sweep or film speed. Probably the best solution is to employ a modified ciné camera possessing a claw or

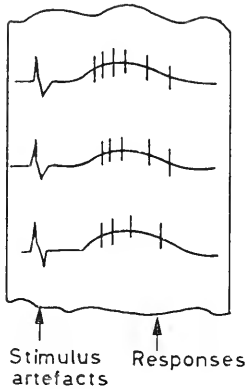


Figure 40.2

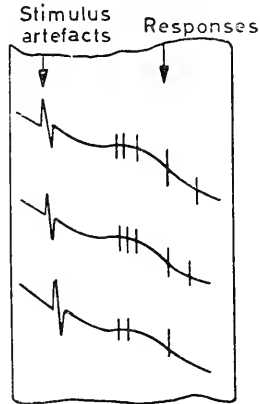


Figure 40.3

maltese cross mechanism and synchronized so that the film is advanced one frame each time the spot flies back, and is motionless during the sweep period.

The mechanical time-base method—For phenomena which are slow, or which cannot be evoked at a definite time, it is better to let the movement of the paper or film itself provide the time-base. The electrical time-base generator is disconnected, and either the cathode ray tube is turned round

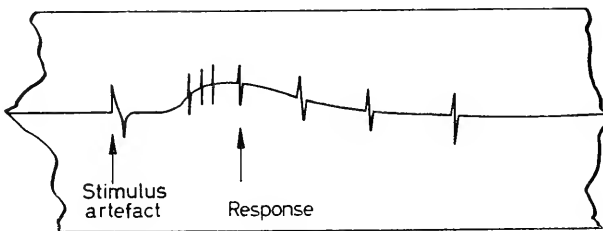


Figure 40.4

bodily on its axis or the deflection amplifier output switched to the *X* plates, so that the signal deflects the spot at right angles to the direction of film motion. The latter method is more convenient but produces in general a change in deflection sensitivity which may be confusing. The record then has perhaps the appearance of *Figure 40.4*.

Film is quite expensive, and not all records that get taken are worthy of retention: conversely, excessively parsimonious use of film may lead to something important being missed. This is the case for the 'intermediate'

METHODS OF RECORDING

recording, which uses a non-consumable medium and which can be edited so that only sections of importance are finally photographed. Magnetic tape is a very suitable and convenient possibility. As conventionally used, tape recorders operate in the audio range, about 50–10,000 c/s, and are therefore satisfactory for storing patterns of action potential discharges. The theory of tape recording has been discussed by Daniel and Axon⁴, and practical requirements for recording machines by Carter⁵. An ordinary commercial tape recorder was used for spike-potentials by Leithead and Thompson⁶, but distortion of the spike on 'playing back' into an oscilloscope is usual. This may be explained in a simplified analysis as follows:

Conventional magnetic recording uses the 'constant current' method, where the recording head causes a magnetization of the tape passing it proportional to the current flowing through its coils and, since it is fed from an amplifier having a high output impedance, proportional to the instantaneous value of the input signal, e . In other words, back e.m.f.'s. occurring across the recording head are swamped and

$$\phi = ke$$

When the tape passes under the replay head an e.m.f. is induced proportional to the rate of change of magnetization

$$e' = k' \frac{d\phi}{dt}$$

Thus if $e = E \sin \omega t$, $e' = kk'\omega E \cos \omega t$, that is, its amplitude rises proportional to frequency for constant E . Thus the signal from a magnetic playback coil rises at 6 dB/octave. In practice it does this up to a limiting frequency at which the whole system breaks down because the playback head gap width becomes comparable with the wavelength of tape magnetization (*Figure 40.5*). This kind of frequency response can be compensated

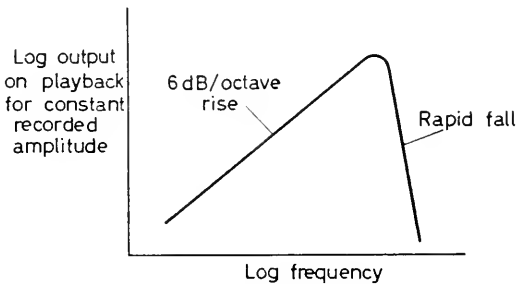


Figure 40.5

for—by suitable filters in the replay amplifier—satisfactorily for audio work, but in commercial machines the compensation is not normally sufficiently exact for the overall frequency (and therefore phase) distortion not to mar the reproduction of waveforms.

To avoid this difficulty Coaton and Whitfield⁷ employed constant-voltage recording, driving the recording head from an amplifier of very low output impedance. Then the back e.m.f. across the head is proportional to the

REFERENCES

instantaneous signal input, and as the back e.m.f. is proportional to $d\phi/dt$ we have on recording

$$\frac{d\phi}{dt} = Ke$$

and on replaying $e' = K' \frac{d\phi}{dt}$

$$\therefore e' = KK'e$$

No compensation is necessary. The authors reported satisfactory reproduction of 1 msec square waves in this manner. The method is not suitable for low frequencies, however, for if $d\phi/dt = Ke$

then
$$\phi = K \int e dt$$

When ϕ reaches tape saturation point the system breaks down and there is therefore a limit to the voltage-time integral for a pulse. Therefore long pulses can only be small.

For recording signals down to zero frequency a system has been described by Daniel⁸ in which constant-current recording is used and on playback a special head is employed which measures ϕ instead of $d\phi/dt$. The manufacture of the necessary head should be within the capabilities of most laboratory workshops, but the author has no experience of the technique and cannot say whether this is in fact so. Recording down to zero frequency can be performed using frequency modulation and a conventional tape deck; since only a little extra electronic circuitry is required this is probably the simpler expedient. An early FM tape recorder was described by Molyneux⁹ and operated in the range 0–100 cycles, satisfactory for cardiographic recording. The method was later extended by the author¹⁰ to work with input signals up to 2 kc/s. This made possible the recording of slow electrophysiological phenomena and, for most purposes, a passable reproduction of spike potentials.

REFERENCES

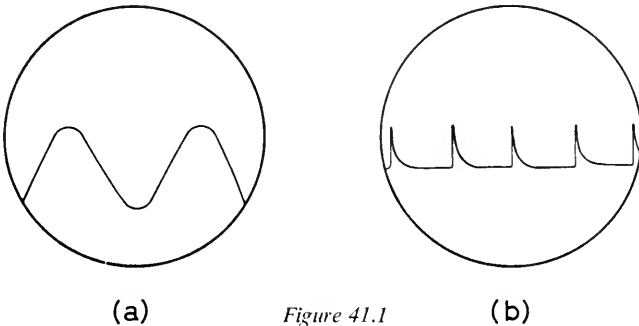
- ¹ Personal communication, 1957
- ² LEVENSON, G. I. P. *J. sci. Instrum.* 27 (1950) 170
- ³ HOLDEN, A. V., LEVENSON, G. I. P. and WELLS, N. *J. sci. Instrum.* 28 (1951) 318
- ⁴ DANIEL, E. D. and AXON, P. E. *J. Instn. elect. Engrs.* Pt. III, 100 (1953) 157
- ⁵ CARTER, J. M. *Wireless World* 59 (1953) 108
- ⁶ LEITHEAD, L. A. and THOMPSON, L. C. *J. Physiol.* 126 (1954) 2P
- ⁷ COATON, J. W. and WHITFIELD, I. C. *J. Physiol.* 125 (1954) 13P
- ⁸ DANIEL, E. D. *Proc. Instn. elect. Engrs.* Pt. B, 102 (1955) 442
- ⁹ MOLYNEUX, L. *Electron. Engng.* 24 (1952) 130
- ¹⁰ DONALDSON, P. E. K. *Electron. Engng.* 27 (1955) 543

TIMERS, COUNTERS AND RATE MEASUREMENT

TIMERS

In protracted experiments in what might be called the 'kymograph' time scale the provision of some form of automatic time marker relieves the worker of the necessity for repeatedly consulting a stopwatch and marking the time of significant events by hand on the smoked paper. In high-speed experiments of the cathode ray tube variety time marks are almost essential; it is usually more satisfactory to record a separate time trace than to rely upon knowledge of the time-base sweep velocity. Even if the latter procedure is adopted a master time standard is still necessary to set the time-base generator up.

Kymographic time markers commonly comprise a synchronous electric clock which drives a bank of cams possessing various numbers of strikers. These operate electrical contacts so that, by appropriate external connection, the circuit is closed at intervals ranging from 1 second to 1 minute. The circuit is completed by a battery and an electromagnetic event marker. The accuracy depends on the degree to which the mains supply frequency approximates to 50 cycles per second. Over a period of days the cumulative error is held at zero, but at any particular instant the frequency may be in error by up to ± 2 per cent. If this is not sufficiently accurate the synchronous clock may be replaced by an electrically maintained pendulum, which rotates a cam system of a similar kind.



Cathode ray tube displays may also employ the 50 cycle mains as a time marker, though if an electrical time base is used the method is not much use for bases shorter than the time of one cycle—20 milliseconds. For longer bases, however, a 50 cycle sine wave is better than nothing.

A difficulty with sine markers is that it is not easy to measure off from the eventual record a distance corresponding exactly to one cycle because the maxima of the wave are so ill-defined (*Figure 41.1a*). A much more satisfactory timer waveform is a train of 'pips' (*Figure 41.1b*). The author

TIMERS, COUNTERS AND RATE MEASUREMENT

has devised a 'Woolworth' timer which converts the 50 cycle mains into a train of pips occurring every 10 msec: only a handful of components is used, as shown in *Figure 41.2*. The waveforms are sketched in *Figure 41.3*.

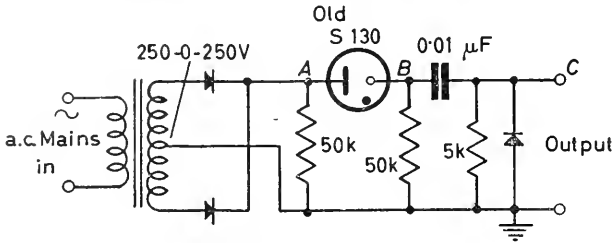


Figure 41.2

At *A* there is 'raw' full-wave rectified a.c. Each time the potential here reached the striking voltage of the regulator tube the tube fires and the potential difference across it drops to the burning voltage, producing a sudden rise of potential at *B*. *B* then follows *A* until the tube extinguishes

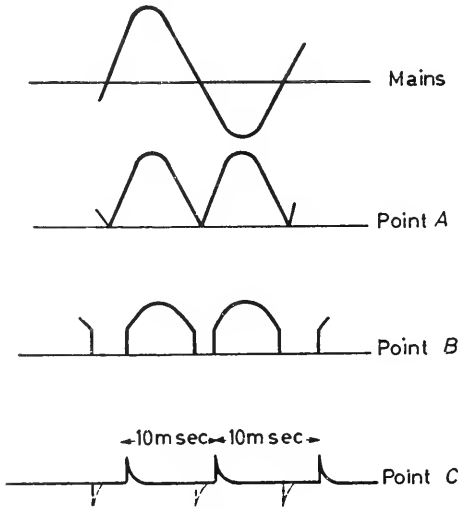


Figure 41.3

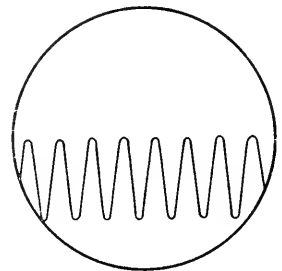


Figure 41.4

again, when it drops back to earth potential. This waveform is differentiated by the $0.01 \mu\text{F}$ capacitor and the 5k resistor so that only the transients are preserved as positive and negative pips; the latter are removed by the crystal diode.

Variable frequency RC oscillators are frequently used for time marking. It is then possible to choose a frequency so that the marker trace has the appearance of *Figure 41.4* rather than *41.1a*. In this case the position of the peaks of the waves is not difficult to measure off. A frequency accuracy of 1 per cent may be expected from the cheaper commercial models, and 0.1 per cent from the superior variety.

For really accurate work recourse may be had to a valve-maintained tuning fork or to a quartz crystal oscillator. The former is capable of

COUNTERS

accuracies of ± 10 parts per million, and the latter of 2 or 3 parts. Tuning fork frequencies are commonly in the high audio range; quartz crystals of the most familiar variety are intended for radio-frequency work and oscillate at upwards of 100 kc/s, but special low-frequency crystals are obtainable which generate oscillations in the range 4–20 kc/s. A single accurately known source of frequency is valuable but not very convenient; usually the utility of the apparatus is extended by using the master frequency to control a chain of frequency-divider circuits, so that timing pips are available with a range of intervals—*Figure 41.5* may be regarded as typical.

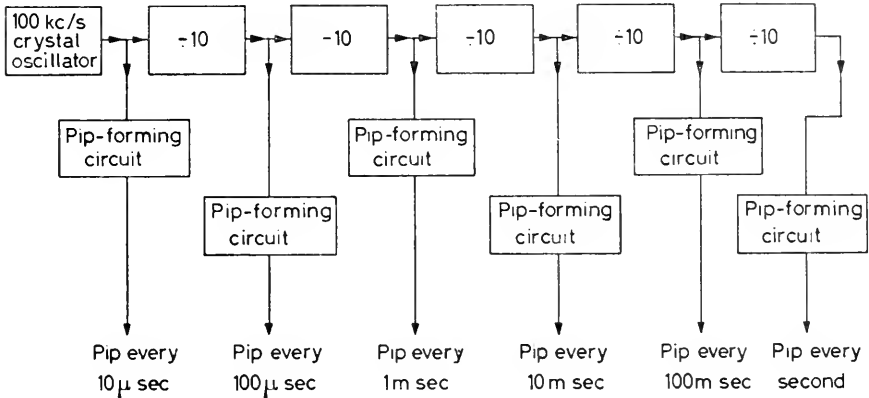


Figure 41.5

Timer units along these lines, using multi-vibrator frequency dividers, have been described by Attew¹ and by Dickinson². The use of phantastron dividers is discussed by Moody and Williams³.

A disadvantage of the frequency-divider circuit is the possibility of the division ratio 'slipping'. For extreme reliability the oscillator may be followed by a chain of 'counter' circuits. As counters are to be the subject of the next section it is not proposed to discuss them here, beyond mentioning that these devices count either on a 'scale of two' or on a 'scale of ten'. The reader will appreciate that whilst an oscillator followed by a $\div 10$ frequency divider, and an oscillator followed by a scale of ten counter, effectively do the same thing, the counter and the divider circuit are quite different. Whilst the latter is an oscillator synchronized to a drive at its tenth harmonic, the former is a circuit which is quite passive until the next drive pulse arrives. Timers employing frequency division may slip and make small mistakes which go unnoticed for a time. Timers employing counters either work properly or make errors so gross that they are recognizable at once. An excellent timer unit employing Dekatron counter circuits and a 10 kc/s crystal has been published by Kay⁴. The author has built this circuit and found it very easy to get going; he has no hesitation in recommending it strongly.

COUNTERS

For low counting rates, up to about 10 per second, the 'Post Office Counter' for registering subscribers' calls at the exchange is suitable and convenient

(Figure 34.28). The total is indicated digitally in a small window up to a maximum of 9,999. These devices are electromagnetic and have coils wound for a variety of impedances; some are only of a few ohms and are not of much use in conjunction with electronic apparatus except via a relay. There is, however, a 500 ohm model which operates on a current of 100 mA, suitable for inclusion in the anode circuit of a small thyratron, and a 2,300 ohm type which is operated by 30 mA, appropriate for control by a power valve.

Drop counters

Drops are counted either conductimetrically or photoelectrically. In the former method drops are used to bridge a pair of electrodes enabling a current to flow which causes, directly or indirectly, the operation of the indicating device. A simple direct system, intended for kymograph marking, is shown in Figure 41.6, in which the indicator is a high-resistance post

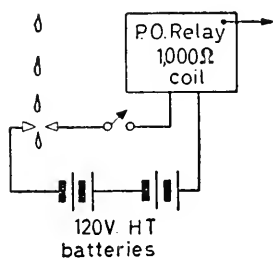


Figure 41.6

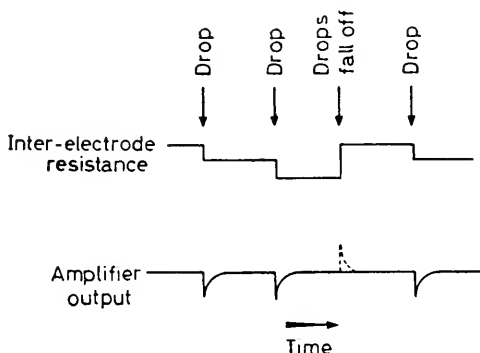


Figure 41.7

office relay with a light pointer attached to the clapper. A similar arrangement, but mains driven, has been demonstrated by Bernstein and Betts⁵. The difficulty with simple counters such as these is the damage done by electrolysis to the body fluid being measured, rendering it unsuitable for recirculation or analysis. In addition, mechanical obstruction arises in the counting cell due to the generation of foam by electrolysed gases. More sophisticated counters employ a valve or thyratron to control the indicator; in this way the current passing through the fluid can be made much smaller.

The mechanical design of drop-counter cells is quite difficult. It must be realized that the device has to accept fluids of a wide range of viscosities and delivery rates. Typically, trouble may be experienced from a drop bridging the electrode gap before its predecessor has fallen off. Campbell and Gilmour⁶ overcame this by arranging the electronic circuit to respond to rate of change of inter-electrode resistance rather than to the resistance itself. They achieved this by using a capacitor-coupled valve amplifier with a time constant shorter than the shortest interval between drops (Figure 41.7). For applications where even the minutest current through the fluid cannot be tolerated the photoelectric drop counter may be used. Hilton and Lywood⁷ may be consulted. These authors report that their apparatus successfully counts drops even of transparent, colourless fluids.

COUNTERS

High-speed counters or 'scalers'

These are important in biological work for applications which include radioactivity measurements and the production of decade time marks from a master oscillator. We begin by considering the hard valve 'scale of two'. Consider the circuit in *Figure 41.8*: it is basically a symmetrical

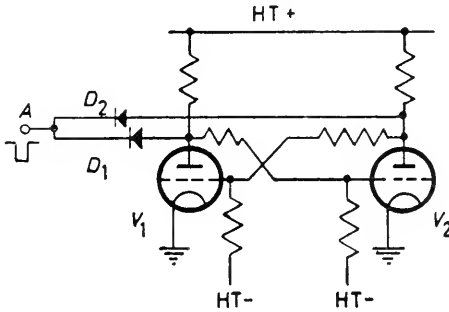


Figure 41.8

Eccles-Jordan circuit with the addition of two diodes. Let point A be normally maintained at $HT+$ potential and let V_2 be conducting and V_1 cut off. Then D_2 will be cut off and D_1 has both terminals at $HT+$. Now let a negative pulse carry A to a potential below $HT+$ but rather higher than V_2 anode. D_2 will remain cut off, but D_1 will now conduct, taking V_1 anode negative and initiating the cumulative effect by which the circuit triggers into the opposite condition— V_1 cut off, V_2 conducting. Because the circuit is symmetrical it follows that a second negative pulse at A will return the circuit to its original condition. Notice that the positive-going back edge of the pulse A has no effect on the circuit, since the only valve it could affect is the one which happens to be conductive, and it is isolated from this because the relevant diode is cut off.

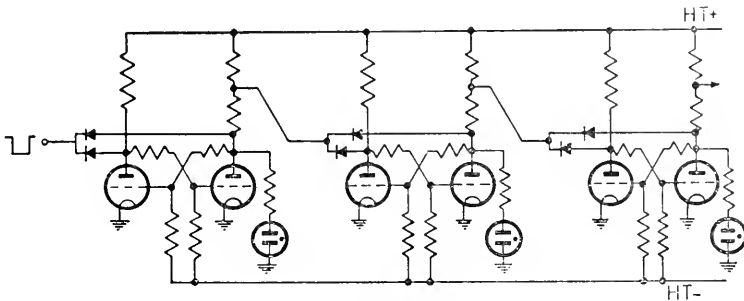


Figure 41.9

Clearly either valve goes through a cycle of operations for every *two* cycles of the driving pulse, and the circuit therefore 'counts down' on a scale of two. It is not difficult to see that the necessary drive conditions are met if A is fed from a tap on the anode load of one of the valves in another, similar, scale of two. In this manner, chains of Eccles-Jordan circuits can be made to count on the scale of two raised to any integral

TIMERS, COUNTERS AND RATE MEASUREMENT

power (*Figure 41.9*); thus a chain of 4 circuits makes a counter on the scale of 16. Most of us are accustomed to count on a scale of ten. Ten is not given by raising two to any integral power and for this reason this circuit is not very convenient as it stands. It may be made to count in decades by a simple modification, as follows:

The presentation of counters of this type is achieved by connecting small neon indicator lamps between the right-hand triode anode of each Eccles-Jordan circuit and earth. When the count is zero all the right-hand triodes are conducting and all the neons are extinguished because the anode potentials are below the necessary maintaining voltage. We call this 'circuits off'. In a scale of 16 the lamp associated with Eccles-Jordan 1 is labelled 1, Eccles-Jordan 2 is labelled 2, Eccles-Jordan 3 is labelled 4 and Eccles-Jordan 4-8, and so on. The total count is the sum of the figures beside the lighted lamps. Thus, a straightforward scale of 16 operates according to the following table:

Count No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	0	1
1st E-J (Ones)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
2nd E-J (Twos)			—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
3rd E-J (Fours)					—	—	—	—	—	—	—	—	—	—	—	—	—	—
4th E-J (Eights)									—	—	—	—	—	—	—	—	—	—

Continuous line = 'Circuit on'

In order to count in decades the cycle has to be modified to this:

Count No.	0	1	2	3	4	5	6	7	8	9	0	1
1st E-J (Ones)		-		-		-		-		-		-
2nd E-J (Twos)			—	—	—	—	—	—	—	—	—	—
3rd E-J (Fours)					—	—	—	—	—	—	—	—
4th E-J (Eights)									—	—	—	—

Comparing the two tables it is clear that to count in decades it is necessary to arrange that on the arrival of the tenth pulse the second Eccles-Jordan is not switched to 'on' by the first, and the fourth circuit is switched to 'off'. This may be achieved by the scheme shown in *Figure 41.10*. The fourth

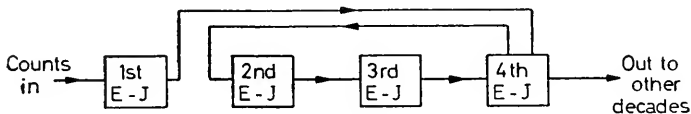


Figure 41.10

Eccles-Jordan circuit is arranged as a 'gate' for pulses passing from the first to the second. The gate is open so long as the fourth circuit is 'off'. Thus the first circuit controls the second in the normal manner until count 10, when it is unable to do so because the gate is closed at count 8; instead, the first circuit output is used to operate the fourth, returning it to 'off' and reopening the gate for the next counting cycle.

COUNTERS

This seemingly complicated behaviour can be achieved quite simply with the help of an extra pair of diodes in a circuit due to Rotblat, Sayle and Thomas⁸. The relevant fragment of circuitry is shown in *Figure 41.11*. When the fourth circuit is 'off', the right-hand triode is conducting, D_2 is cut off and pulses are transferred from the fourth to the second Eccles-Jordan circuit via D_1 . When the fourth circuit is 'on', the right-hand anode rises to HT+, rendering D_2 conductive. At the same time, the left-hand anode is now negative to HT+, D_3 conducts and cuts off D_1 . In this manner pulses arising from Eccles-Jordan No. 1 are directed to Eccles-Jordan No. 4 and blocked from Eccles-Jordan No. 2 (for further details the original

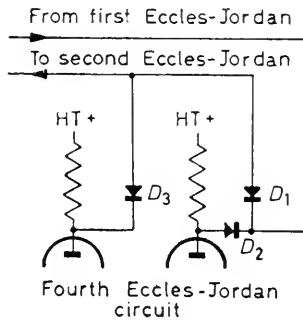


Figure 41.11

article should be consulted). Scale of ten counters can be connected in chains to form scales of 100, or 1,000 and so on. A cheaper kind of scaler, using Eccles-Jordan circuits but relying on capacitors instead of diodes for coupling, has been discussed by Gottlieb⁹. It seems to the author that circuits of this kind might be rather more difficult to get going than the diode-coupled variety. Eccles-Jordan design procedure has been considered by Taub¹⁰.

Scalers of the Eccles-Jordan type are reliable and widely used, but they contain a large number of valves, and special decade counting tubes have been developed to effect simplification and gain economy here. These have been reviewed by Kandiah¹¹. Briefly, they fall into two categories:

(1) *Glow discharge types* in which a glow occupies one of ten positions, circularly arranged, in a soft valve. Each input pulse moves the glow on to the next station, and the count is read off by observing the position of the glow against an external annular numbered escutcheon. At the completion of each complete circuit of the glow the tube delivers a 'carry' pulse which may be used to operate further tubes.

(2) *Cathode ray types*, in which a cathode ray beam in a hard valve is arranged to occupy one of ten stable positions. At each position the electron beam bombards an appropriate numeral, made of fluorescent paint, causing it to glow and enabling the count to be read off. The beam is transferred to the next position by setting up electric fields within the tube with suitable electrodes.

Glow discharge decade counter tubes are known variously as Dekatrons or Nomotrons, according to the manufacturer. Cathode ray types include

TIMERS, COUNTERS AND RATE MEASUREMENT

the Trochotron and the Philips E1 T; they have the advantage of speed, in that count rates of the order of 200,000 per second are possible. The fastest glow tubes can count at about a tenth of this rate.

For biological work the glow-discharge types are usually fast enough and it is not proposed to consider cathode ray counters further. As examples of the former, the Ericsson Dekatrons type GC10 B and GC10 D will be briefly discussed; the only *operational* difference between these two tubes is their upper limiting counting speeds, 4,000 counts per second with the former, 20,000 counts per second with the latter.

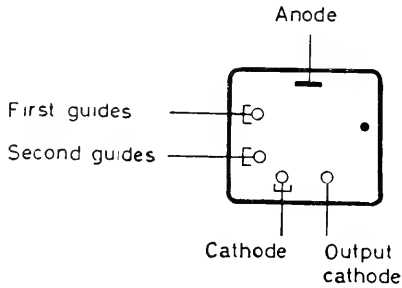


Figure 41.12

The GC 10 B—The GC 10 B comprises a central disc-shaped anode surrounded peripherally by 30 equally spaced rod-like electrodes. Electrodes 1-4-7-10...25 are connected together internally and brought to an external connection as 'cathodes'. Electrode 28 is brought out separately

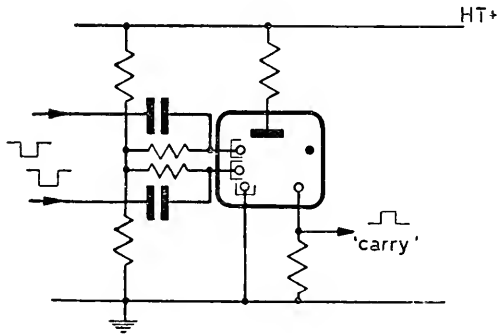


Figure 41.13

as the 'output' or 'carry' cathode. Electrodes 2-5-8...29 are connected together and brought out as 'first guides' and similarly 3-6-9...30 as 'second guides'. The gas filling is neon. The circuit diagram is as in Figure 41.12. Consider now the circuit in Figure 41.13. On applying the HT the glow strikes between the anode and one of the cathodes, the n th say. There is no possibility of the glow bridging the anode and one of the guides because the latter are biased such that the potential difference is insufficient. However, if a negative pulse of sufficient magnitude be applied via a

COUNTERS

capacitor to the first guide, the glow transfers one electrode round the ring: if a fraction of a second later a similar negative pulse is then applied to the second guide, the glow advances a step further. At the conclusion of the second pulse, the glow is nearer to the $n + 1$ th cathode than to the n th, and in consequence moves forward once more, thus accomplishing one count. At the completion of each ten counts the glow passes the output cathode, and a positive pulse is developed across the load resistor by the passage of the glow current through it.

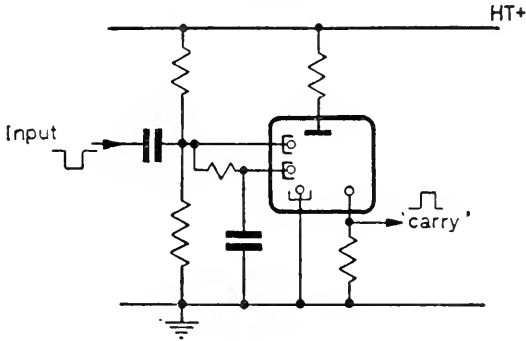


Figure 41.14

The GC 10 B is, for obvious reasons, called a 'double pulse' dekatron, and driving it is seen to be a matter of providing two negative pulses in quick succession. This may be done in a number of ways; in most of them a negative pulse is applied directly to the first guides and via a low-pass filter to the second, as in Figure 41.14. With properly chosen values the waveforms are then as sketched in Figure 41.15.

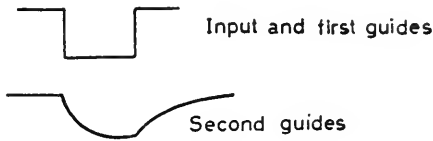


Figure 41.15

The output pulse from one dekatron is in the wrong phase and of insufficient magnitude to drive another. A circuit recommended by the manufacturers which effects the necessary phase reversal, amplification and double pulse production is shown in Figure 41.16. It makes use of the high-speed trigger-tetrode type GTE 175M, which is convenient in that no cathode heater power is required. This circuit works up to 500 counts/sec.

The GC 10 D—This high-speed dekatron has 40 electrodes equally spaced around the anode, made up of 10 cathodes and 3 sets of guides. All first guides are brought out together as one connection, and all second guides as another. The third guides are divided, as are the cathodes, into output third guide and all other third guides, and output cathode and all other cathodes: the output third guide, of course, precedes the output cathode. The gas filling appears to be argon. The circuit diagram is as in

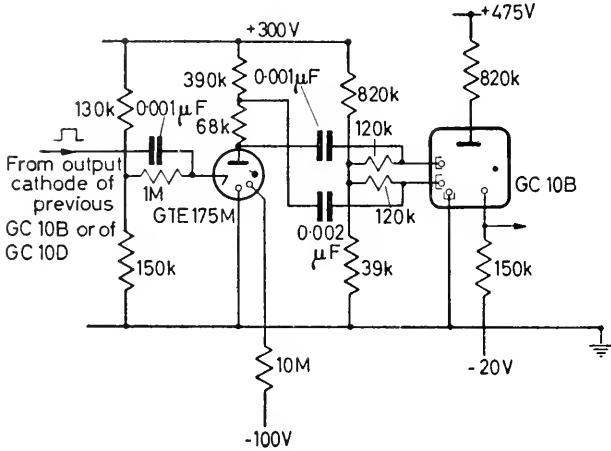


Figure 41.16

Figure 41.17. Glow transfer is achieved in the GC 10 D by applying a single pulse only. This is done by making use of both the leading and trailing edges of the transfer pulse, and by the phenomenon called 'auto transfer', which may be explained as follows:

In Figure 41.18 we have the anode and two electrodes *A* and *B*, which may be cathodes or guides. If the glow can by some means be caused to alight on electrode *A*, the glow current will charge *C* to a definite potential difference determined by *R*. If by this process *A* becomes sufficiently positive the glow 'auto-transfers' to the more negative electrode *B*. *C* then discharges again through *R*, but the glow will remain at *B*.

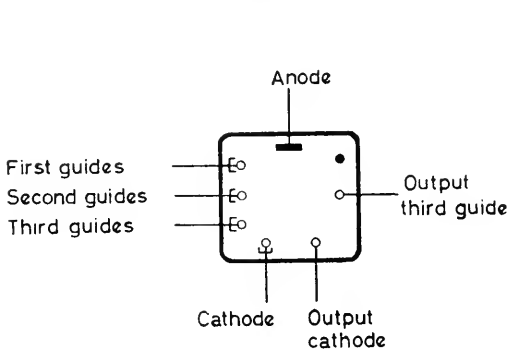


Figure 41.17

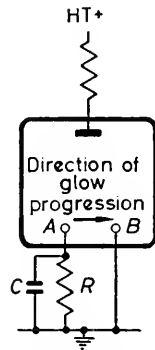


Figure 41.18

Thus a typical GC 10 D drive circuit is shown in Figure 41.19. Let the glow be on the *n*th cathode. On the arrival of the leading edge of the transfer pulse the first and second guides are driven sufficiently negative to draw the glow on to the first guide, whereupon it auto-transfers to the second. The back edge of the transfer pulse then comes along, driving the first and second guides positive again. The third guides are now relatively

COUNTERS

very negative and in consequence the glow moves there, only to be auto-transferred to the $n + 1$ th cathode.

It might seem at first sight that the first and third guides might be dispensed with; that the leading edge of the transfer pulse could be used to move the glow from n th cathode to a guide, and the trailing edge would then move it on the $n + 1$ th cathode. The difficulty is to prevent its moving back to the n th again. In point of fact single-guide tubes do exist; specially shaped electrodes ensure that transfer occurs in the required direction.

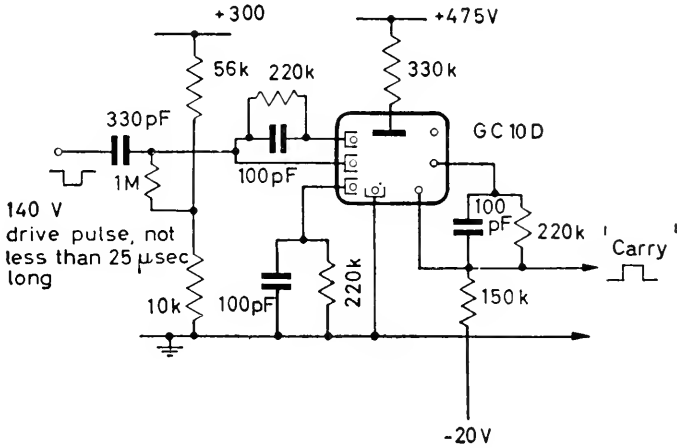


Figure 41.19

A GC 10 D can be driven from another GC 10 D via a GTE 175 M high-speed trigger tetrode at rates up to 1,000/sec. The necessary circuit is shown in Figure 41.20, and is reproduced from part of reference 4 in this chapter.

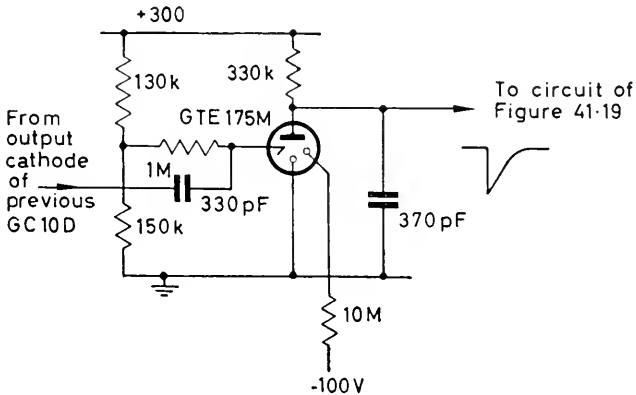


Figure 41.20

A scaler for radioactivity measurements using dekatron counters has been described by Florida and Williamson¹². Kerkut¹³ has given details of a dekatron action-potential counter. The time marker, employing dekatrons, published by Kay has already been mentioned.

RATE MEASUREMENT

It is often the frequency with which events happen, rather than the events themselves, which it is convenient for apparatus to measure. In this section some rate-measuring devices are considered.

For rates up to 10/sec an electromagnetic device such as the Thorp counter is satisfactory. In this apparatus each event—for example the falling of a drop—causes a magnetically operated pawl to raise a toothed rack one notch. The toothed rack carries a writing point which marks a smoked drum or similar recording medium. At regular intervals an impulse from

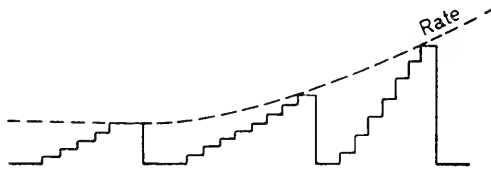


Figure 41.21

a timer, usually of the synchronous electric clock type, retracts the pawl, allowing the rack to fall to the bottom of its range of travel. The record as presented has the appearance of *Figure 41.21*; the heights of the 'staircases' give a measure of the rate.

For higher rates of counting electronic methods can be employed. The simplest kind of event-rate measuring device is a pulse generator feeding a single-stage RC low-pass filter. Typically, what has to be measured is a train of voltage 'pips' of no particular size or duration. Suppose these pips are used to trigger a flip-flop, producing pulses of known amplitude and length, and that these flip-flop pulses are fed into an RC circuit (*Figure 41.22*). Suppose further that, for the moment, the pips are arriving at

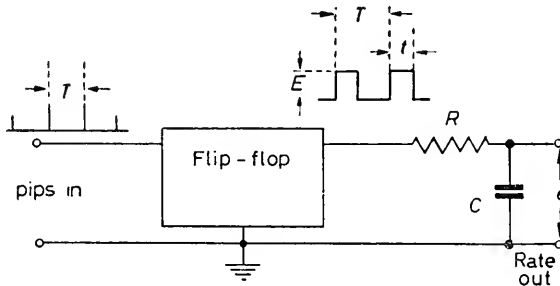


Figure 41.22

constant intervals T . Let the flip-flop have constant-voltage output characteristics and let it generate pulses of amplitude E and of duration t . Then the final voltage e to which C charges is given merely by

$$\frac{e}{E} = \frac{t}{T - t}$$

REFERENCES

Since the rate = $1/T$, it follows that provided $t \ll T$

$$e = E \times t \times \text{rate}$$

The system is open to two objections. One is the requirement that t be much smaller than T , which means the arrangement becomes non-linear at high count rates. The other is that when the input rate changes, e responds only with a time constant RC , and if any attempt is made to shorten the response time by reducing RC , the output voltage then fluctuates seriously in step with each operation of the flip-flop. Much can be done to remove these fluctuations by subsequent filtering but slow response remains as a criticism and is in fact fundamental to this simple rate-measuring scheme. Nevertheless, in this or in more elaborate forms, it is often used, e.g. by Boyd and Eadie¹⁴ in a heart-rate meter and in the author's frequency-modulation tape recorder¹⁵, where such a circuit is employed to measure the frequency of flux-reversal in the magnetic tape.

In an effort to produce a circuit which responds instantly to a rate change, Andrew and Roberts¹⁶ and Manzotti¹⁷ have devised apparatus which computes the reciprocal of the interval between the n th event and the $n + 1$ th, presenting the answer immediately upon arrival of the latter. This is of course as quick-acting as is possible. The former authors' apparatus is for displaying the rate of discharge of action potentials; Manzotti's is a cardiometer. The way in which the requisite calculation is done in these devices makes interesting reading, and it is not proposed to discuss them here. The writer would, however, beg any reader contemplating quick-acting tachometry to consider carefully whether the technique is really necessary, since the apparatus required is liable to be complex. The cardiometer mentioned employs 11 valves, most of which are double, and the nerve discharge-rate meter contains 27.

REFERENCES

- ¹ ATTEW, J. E. *Wireless World* 58 (1952) 114
- ² DICKINSON, C. J. *Electrophysiological Technique*, Electronic Engineering
- ³ MOODY, N. F. and WILLIAMS, F. C. *J. Instn. elect. Engrs.* Pt. IIIA, 93 (1946) 1188
- ⁴ KAY, R. H. *Electron. Engng.* 28 (1956) 452
- ⁵ BERNSTEIN, L. and BETTS, J. C. *J. Physiol.* 110 (1949) 1P
- ⁶ CAMPBELL, F. W. and GILMOUR, W. G. *J. Physiol.* 111 (1950) 2P
- ⁷ HILTON, S. M. and LYWOOD, D. W. *J. Physiol.* 123 (1954) 64P
- ⁸ ROTBLAT, J., SAYLE, E. A. and THOMAS, D. G. A. *J. sci. Instrum.* 25 (1948) 33
- ⁹ GOTTLIEB, I. *Wireless World* 60 (1954) 234
- ¹⁰ TAUB, D. M. *Electron. Engng.* 27 (1955) 386
- ¹¹ KANDIAH, K. *Electron. Engng.* 26 (1954) 56
- ¹² FLORIDA, C. D. and WILLIAMSON, R. *Electron. Engng.* 26 (1954) 186
- ¹³ KERKUT, G. A. *Electron. Engng.* 27 (1955) 378
- ¹⁴ BOYD, W. E. and EADIE, W. R. *Electron. Engng.* 26 (1954) 330
- ¹⁵ DONALDSON, P. E. K. *Electron. Engng.* 27 (1955) 543
- ¹⁶ ANDREW, A. M. and ROBERTS, T. D. M. *Electron. Engng.* 26 (1954) 469
- ¹⁷ MANZOTTI, M. *Electron. Engng.* 28 (1956) 446

LAYOUT AND THE CONTROL OF INTERFERENCE

The interference to be discussed in this chapter is that due to the action of stray fields on the input circuit of high-gain amplifiers. Spurious responses generated by the apparatus itself, e.g. due to intermittent faults, inadequately smoothed or stabilized power supplies, etc., are dealt with in Chapter 43. Interference minimization will be considered with special reference to electrophysiology, where the problem is particularly acute, but the principles of interference control are quite general and applicable in other work: these principles should be understood because they have to be applied whenever a new experiment is set up. The apparatus designer can, by appropriate design, do something towards producing gear which is insensitive to certain

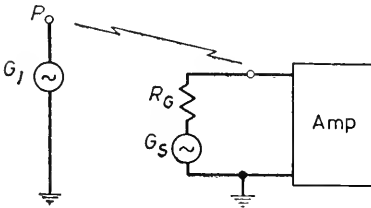


Figure 42.1

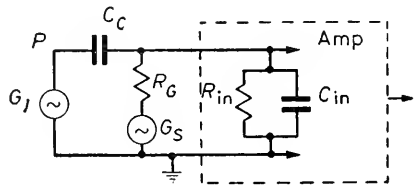


Figure 42.2

kinds of interference, e.g. by using differential amplification, but it is chiefly with the user that the onus of obtaining unadulterated signals at the amplifier input rests.

Interference is borne in three ways:

Electric fields—Most interfering electric fields arise in the manner shown in Figure 42.1; the point P , driven by the interference generator G_I , alternates in potential with respect to earth in the neighbourhood of the amplifier input terminals. The equivalent circuit is shown in Figure 42.2. C_c is the coupling capacitance between P and the 'live' input lead, C_{in} the amplifier input capacitance, R_{in} the input resistance, and R_G is the signal generator resistance—in electrophysiology an electrode.

Clearly the effect produced by P is minimal when C_c is small (which is simply saying that P should be removed as far as possible) or when C_{in} is large, or when R_G or R_{in} are small. Since the efficiency of signal transfer of the arrangement is poor unless R_{in} is much greater than R_G , it follows that in practice it is R_G rather than R_{in} which will be the determining factor.

It follows that experiments using silver wire or wick electrodes are easier to carry out than those employing microelectrodes; for the former we have R_G of the order of ohms $\times 10^4$, whereas for the latter the figure is about a thousand times greater. Furthermore, with the low resistance electrode, C_{in} may be allowed to be quite large, perhaps 100 pF, but the same value in

conjunction with a microelectrode would produce an excessive high-frequency loss.

A powerful weapon against interference of this type is the electrostatic screen, a conductor interposed between source and amplifier, and earthed

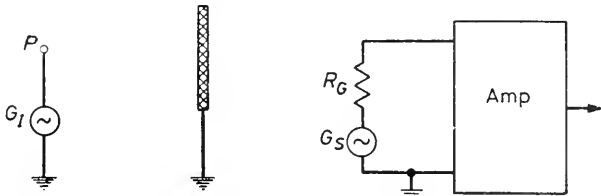


Figure 42.3

(Figure 42.3). Lines of electric force emanating from P in the direction of the amplifier now alight on the screen and, even though it be of mesh construction, are prevented from reaching the amplifier. The equivalent circuit is Figure 42.4. Notice that the earthing of the screen must be sound, other-

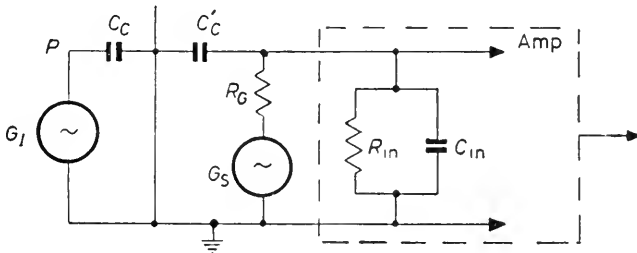


Figure 42.4

wise the equivalent circuit degenerates to Figure 42.5, we have a high-pass filter, and interference proceeds as before.

Magnetic fields—According to the laws of electromagnetic induction a steady magnetic field cannot cause interference but a fluctuating one can,

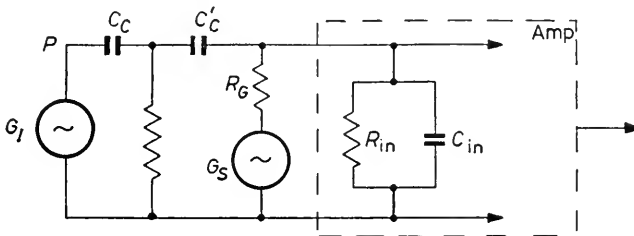


Figure 42.5

and in practice such fields are set up by alternating currents flowing through devices possessing solenoidal windings such as transformers, heaters, and electric motors, and to a lesser extent by apparatus possessing what is virtually a single turn winding, notably electric light bulbs. The magnetic field so produced threads the single turn made by the input circuit of the amplifier, Figure 42.6, inducing e.m.f. into it. The magnitude of this e.m.f.,

LAYOUT AND THE CONTROL OF INTERFERENCE

for a given field, is proportional to the area of the single turn or 'loop'. Thus *Figure 42.7a* is a bad arrangement but *b* is good. The interfering e.m.f. is also proportional to the sine of the angle between the plane of the loop and the direction of the magnetic lines of force. If the loop is already small it is

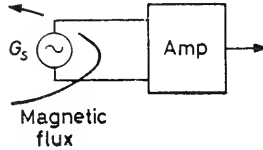


Figure 42.6

difficult to minimize the interference by altering the plane of the loop. Fortunately it is often possible to alter the direction of the field.

Radiated electromagnetic fields—Whenever an alternating electric current flows round a circuit, systems of electric and magnetic fields may be thought of as growing out of, and collapsing back into, the conductors in sympathy with the successive growth and decay of current. These fields represent energy alternatively derived from, and returned to, the electric circuit. However, not all the energy associated with each growth is returned in the subsequent collapse. A fraction of it is 'radiated' off into space as an electromagnetic or 'radio' wave. The effect is extremely small at 50 c/s but becomes important as the frequency is raised into the hundreds-of-kilocycles region,

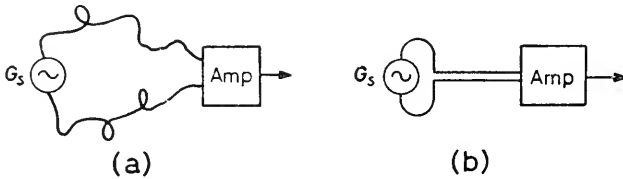


Figure 42.7

and above. An electromagnetic wave, as the name suggests, contains both electric and magnetic field components and cannot exist unless both are present. It follows that any system of protection against either pure electric field- or pure magnetic field-borne interference will be effective against radiation, and no specific remedy is necessary.

Causes of interference

The most ubiquitous kind of interference is due to the supply mains, which shows up as the all-too familiar 50 cycles wavy line on cathode ray tube traces, and as 'hum' on loudspeakers. It may be borne by electric or magnetic fields, often by both. Another variety is 'impulsive' interference, which appears as a sharp spike and is heard as a 'click'. It is usually electric-field-borne and is caused by the sudden appearance of a large back-e.m.f. when a mains circuit is broken which contains an inductive load. Switches, thermostats, commutator motors, etc., are potential sources of trouble here.

The unexpected reception of the B.B.C. is, of course, interference from radiation. For a biological amplifier to operate as a radio set it is necessary that there be a rectification process somewhere in the input circuit, but this

is probably the rule rather than the exception. A small amount of non-linearity in the amplifier will suffice, but even without this, in a circuit containing brass stereotaxic apparatus, animal preparation, silver wires, agar bridges, etc. it is hardly surprising if a rectifying junction exists. It is not sufficient to be remote from a B.B.C. transmitter to avoid interference of this kind. Most laboratories are near towns large enough to support radio-controlled taxis, police cars, road patrols, and in addition there are uses for powerful radio-frequency sources other than for communications, e.g., dielectric heaters and induction furnaces.

Prevention of interference

Preventive techniques for interference fall under three headings; those which seek to cure it at its source, those which protect the preparation, and those which protect the amplifier.

Prevention at source—electric field—For minimizing 50 cycle pick-up, the golden rule is to screen everything possible which is fed from the mains. The laboratory wiring should be in lead-covered cable or in steel conduit, and power points should be of the 3 pin metal-clad variety. All flexible leads should be of screened cable, possessing an outer protective PVC sheath* with the braiding connected to the earthing pin on the power plug. All the appliances fed by these cables should as far as possible have outer cases of metal—rather than plastic or some other insulating material—and the cases earthed via the braiding on the cable. Bench lights should have metal, not bakelite, lamp holders, again properly earthed, supporting a spun-metal lampshade. The reader will be able to continue in this vein for himself.

The rule that, if an appliance is to be switched in one pole only then the switch must be connected in the phase, not the neutral pole, is sound engineering practice and is made in the interests of safety. It is also extremely important in interference reduction, since the phase side of the mains is the one which causes all the trouble; the neutral lead is very nearly at earth potential. By switching off appliances at the power socket in the *phase* side the whole appliance is then earthed. If the switch is in the neutral lead, the appliance continues to cause interference—though the screening will do much to reduce it—besides being unsafe.

The electric fields responsible for impulsive interference are contained automatically by the techniques which are applicable to 50 cycle interference prevention. As an additional safeguard it is as well to reduce the back-e.m.f.'s which cause them, and this can be done by connecting across offending contacts the simple series CR suppressor circuit of *Figure 42.8*. The best component values should be found by trial and error; those shown in the figure are suitable for a first attempt. The circuit may be effective across the brushes of a commutator motor, but if possible it is better to change over to the induction variety, if necessary with variable speed gearbox.

Prevention at source—magnetic fields—The most reliable ally here is distance, and every effort should be made to keep gear carrying alternating current as far away from the amplifier input circuit as possible—say, at least 6 feet. If a bright light is needed immediately over the preparation it

* Exposed braiding collects dust.

can be supplied by a motor-car headlamp bulb fed from an accumulator (but *not* the same accumulator as that used to heat direct-coupled pre-amplifier valves) and mammalian or avian preparations can be kept warm by low voltage heaters similarly fed. If the near presence of some mains-driven items is absolutely necessary it may be feasible to reduce the interference caused by it by reorienting it so that its field cuts the pick-up loop at a shallower angle. If there is any freedom of choice in its construction, 'astatic' winding may be possible; the winding being split into two halves, arranged so that the fields set up by the two parts oppose one another and cancel (Figure 42.9).

Finally, of course, all mains-driven apparatus must be supplied via the usual twin cable and not by two distinct insulated wires. With the twin

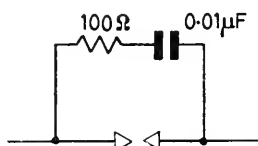


Figure 42.8

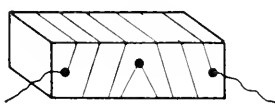


Figure 42.9. One method of astatically winding a heater element

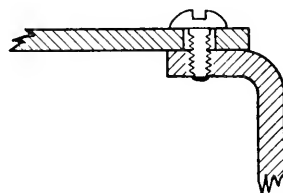


Figure 42.10

cable the two conductors lie close together and the magnetic fields due to each largely cancel one another. If the two conductors are allowed to separate an interfering magnetic field is created proportional to the area of the loop so formed.

Prevention at source—radiation—Unless one is a person of considerable influence there is very little which can be done about this. So far as mobile radio users are concerned, the messages passed are usually brief, and the research worker need only exercise a little patience.

Amplifier protection—electric fields—The use of aluminium or steel cases for electronic gear is now general and is usually sufficient to screen adequately the internal input wiring of the amplifier. The only point to watch is that the pieces of sheet metal composing the cases are not isolated electrically from one another by paint; Figure 42.10 shows two such pieces joined together by a self-tapping screw. The hole in the lower member is drilled such that the screw thread grips it tightly, but the hole in the upper member is sufficiently large to allow the thread to pass through easily. Under these circumstances there is little risk that there will not be good electrical connection from screw to lower member, but that from screw to upper may be very poor. It is usually recommended that the paint be scraped away from the region under the screw head to cure this, but the procedure is liable to produce messy looking work. A better practice is to fit an ordinary cheese-headed bolt into the clearance hole before painting, securing it with a nut. After painting, the screw is removed, leaving a neat circular area of bare metal.

Outside the body of the amplifier the input leads should be of screened cable, polythene insulated if a low capacitance is necessary, and if cathode follower probes are used, the leads from microelectrode to cathode followers

should be double- (i.e. cathodally) screened, polythene insulated, in the interests of good high-frequency response.

Amplifier protection—magnetic fields—Whilst the greater part of this section is concerned with minimizing the area of the loop formed by the input circuit, this is an appropriate point to draw attention to the fact that some kinds of amplifier, e.g. tape recorders and microphone amplifiers, possess an input transformer to match the playback head or microphone to the first valve grid. The secondary winding of such a transformer consists of an extremely large number of turns, each of which represents a small loop, and is in consequence very prone to hum pick-up from magnetic fields. The

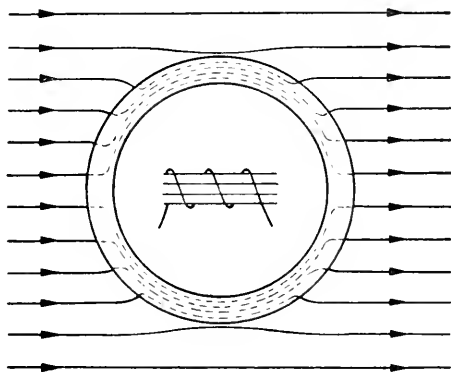


Figure 42.11

solution is to enclose it in a mumetal box. The highly permeable mumetal provides a by-pass for the magnetic flux, as indicated in *Figure 42.11*. The magnetic properties of this useful material are destroyed by any sort of cold mechanical working, and a special heat treatment is necessary to recover them. For this reason the fabrication of such boxes is best left to professionals; most manufacturers of signal transformers can supply suitable boxes for their products.

With biological amplifiers the mechanical layout of the first amplifying stage is all-important. If the amplifier is double-sided, the valves should be placed close together, and the leads joining them to the electrodes, whether via cathode follower probes or not, must run close together (*Figure 42.12*). In addition the preparation earth should be made by a short wire to the screening of one of the input cables. In the case of single-sided amplifiers, the braiding of the screened cable must be used to provide the return path for the signal current after passage through the input circuit of the amplifier, and this braiding must be earthed at the amplifier *at one point only*, namely the earthy end of the first stage cathode biasing resistor (*Figure 42.13*). The screening must *not* make electrical contact with the amplifier case where it enters it. If a single pole connector is used of the popular coaxial type in which the fixing lugs are electrically continuous with the outer conductor, then these lugs must be insulated from the case by suitable fibre washers. If an input potentiometer is included it must be wired as shown in *Figure 42.14*.

Preparation protection—electric fields—In any form of electrophysiological

LAYOUT AND THE CONTROL OF INTERFERENCE

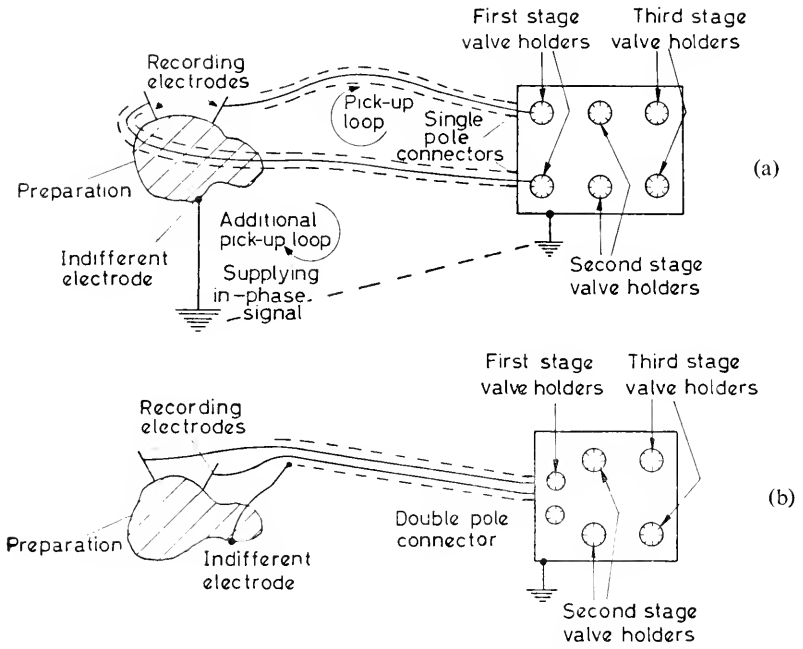


Figure 42.12 (a) Bad arrangement; (b) good arrangement

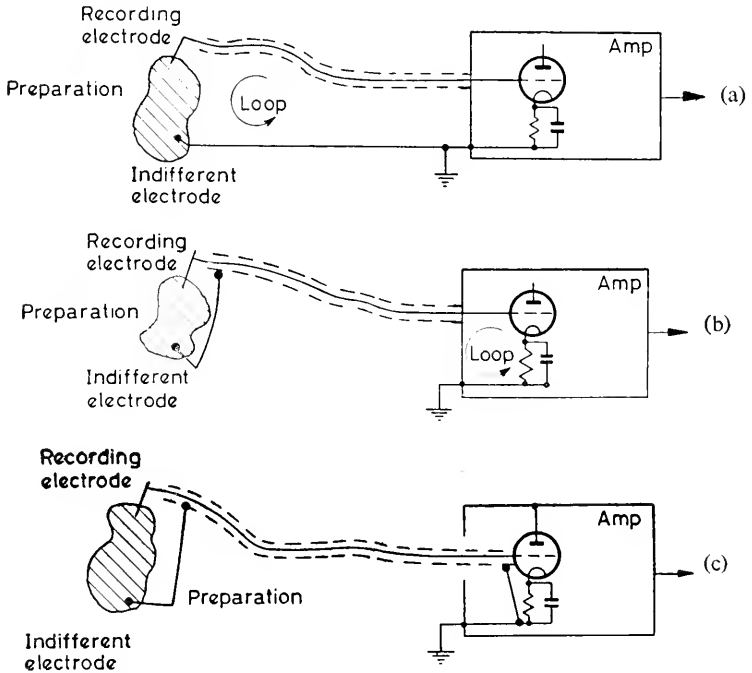


Figure 42.13 (a) Bad; (b) rather bad; (c) good arrangement



Plate 42.1

LAYOUT AND THE CONTROL OF INTERFERENCE

work involving human subjects, e.g. cardiography and encephalography, the imprisonment of the patient in some form of wire screening cage is clearly undesirable and it is fortunate that it is also usually unnecessary. In cardiography the electrodes are of extremely low resistance and therefore not subject to serious electric interference; furthermore the signal is relatively

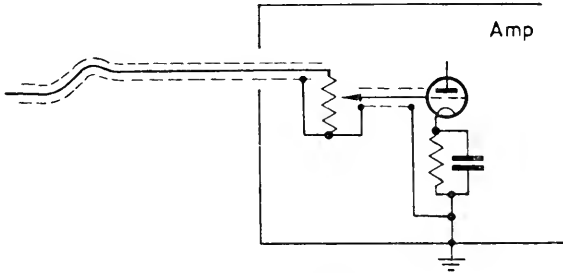


Figure 42.14

large (2 or 3 mV). In encephalography the electrodes are also of low resistance but the signal is much smaller, of the order of μV . Fortunately the electrodes on the patient's scalp are quite close together and subject to similar interference effects. Thus full advantage can be taken of the properties of differential amplification.

In research work on excised preparations, however, electrode resistances are generally much higher and some form of screening cage is frequently essential. The precise form which this takes depends on the size of the animal and on personal taste. In the Cambridge School of Physiology under Matthews the cage is of sufficient size to contain animal, workbench, experimenter and battery driven pre-amplifier (*Plate 42.1*). At the National Institute for Medical Research, Mill Hill, the animal rests on an earthed metal table over which is erected a sort of screening gantry. This provides screening at top, bottom and the two ends, leaving two sides for access by the experimenter. The sides may be closed, if necessary, by clipping on special panels. The cathode-follower probe units which feed the recording amplifier, and the RF stimulating probe units, are mounted inside the gantry; the whole of the rest of the apparatus is carried on 19 in. racks in a different part of the room.

An open-mesh structure, say $\frac{1}{2}$ in. wire netting, suitably braced, is quite satisfactory for such cages, but care must be taken, if a number of separate pieces are used, to ensure electrical continuity between them, preferably by soldering.

Preparation protection—magnetic fields—The rule here is simple. Keep all single core leads as short as possible and all recording electrodes as close together as the terms of the experiment and working convenience allow.

Identifying nature of interfering fields—A quick method of detecting whether 50 cycle pick-up is electrically or magnetically borne is to make use of the fact that electric pick-up is a function of electrode resistance and magnetic pick-up is usually not. Thus irrigation with Ringer's fluid of an excised nerve lying over a silver-wire recording electrode will reduce the amplitude of the pick-up if it is electric, but not if it is magnetic. If the

LAYOUT AND THE CONTROL OF INTERFERENCE

input resistance of the amplifier is low, such irrigation may actually increase magnetic pick-up.

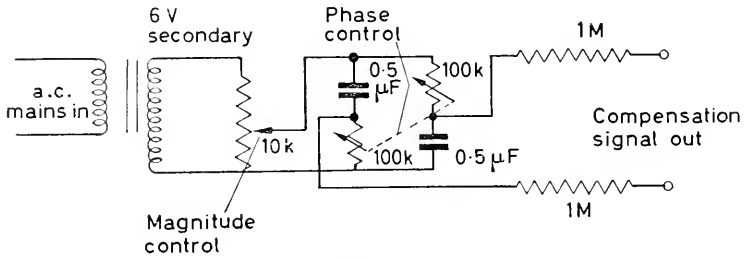


Figure 42.15

Compensation for 50 cycle pick-up

In specially intractable cases, where all else has failed, it may be worthwhile to introduce into the amplifier at some suitable point a compensating 50 cycle signal, variable in amplitude and in phase, from some such device as that outlined in *Figure 42.15*.

FAULTS

The number of types of fault which commonly overcome individual components is not great, but the list of symptoms which these faults can produce in complex electronic equipment is endless. The most that can be undertaken in this chapter is to list the more important component failures, and to indicate the kinds of diagnostic procedure to adopt when complicated units break down.

COMPONENT FAULTS

These fall into two categories; failures which occur suddenly, and those which are progressive. The fact that some components are more subject to one type than to the other is of considerable value in the diagnosis of equipment faults, since a gradually developing component failure is usually mirrored by steadily deteriorating equipment performance. Common component faults may be tabulated as shown on page 666.

EQUIPMENT FAULTS

We consider now faults within a complete unit of apparatus. We may conveniently divide these into:

- (1) complete failure, in which no part of the apparatus works
- (2) partial failure, in which part of the apparatus works
- (3) intermittent failure, in which all the apparatus works some of the time
- (4) illusory failure, in which it is in fact the operating procedure which is at fault.

(1) *Complete failures*

Complete failures are those in which, on switching on, either nothing happens or else the fuses blow instantly, and do so again when replaced. In-so-far as in most electronic circuits the only thing which is common to the whole equipment is the power supply, attention should first be directed towards the power pack and HT divider circuits. If nothing happens on switching on, see if the valve heaters are lighting up: if they are not, check the mains transformer primary circuit; if they are, the trouble is probably absence of HT voltage—find out if this is so. If experiment confirms this, check the alternating voltage across the transformer secondary winding, then find out if there is an appropriate direct voltage across the rectifier input capacitor. Follow the d.v. through the smoothing choke. Advance by steps in like manner in logical sequence until the cause of the trouble is found, perhaps an open-circuited choke. Unless the apparatus has been built recently by oneself and one's memory is good, it is almost essential to have a circuit diagram at hand; given this, the procedure is perfectly simple and straightforward.

The apparatus which blows its fuses is rather less easy because one has no

FAULTS

<i>Item</i>	<i>Progressive</i>	<i>Sudden</i>
Composition resistors	Drift in value due to temperature cycling	Rare
Wirewound resistors	Intermittent resistance changes due to fractured welds	Open circuit, due to burn-out of wire
Pyrolytic resistors		Open-circuit due to burn-out of spiral track
Capacitors	Sporadic breakdown of insulation	Short circuit—complete failure of insulation. Occasional Open circuit—due to a connecting wire pulling away from its metal foil
Electrolytic capacitors	Loss of capacitance and increase of leakage current with old age	Leakage current becomes so great that the temperature rise produced leads to evolution of gas and eventual explosion. (This should never happen if the circuit is properly fused)
Potentiometers	With old age, slider begins to make poor contact	Rare
EHT transformers	Gradual failure of insulation	Winding open circuit—usually a burn-out as a result of overload by a fault condition elsewhere (e.g. electrolytic capacitors)
HT transformers	Rare	
Valves	Gradual loss of cathode emission	Heater open-circuit
Regulator tubes	Gradual drift of running voltage	Rare
Wiring	Corroded soldered joints due to use of unsatisfactory flux. 'Dry' joints, due to inefficient cleaning of surfaces before soldering	Short circuits caused by the bridging of soldering tags by metal fragments (e.g. drilling chips) inadvertently left behind after assembly of work

EQUIPMENT FAULTS

time in which to make the necessary checks. The correct approach here is by a kind of graduated amputation. With a good supply of spare fuses available, find the positive terminal of the HT smoothing capacitor and unsolder the HT+ busbar supplying the whole of the rest of the apparatus. This isolates the power pack. If the fuses again blow upon switching on, suspect the smoothing capacitor or the rectifier input capacitor of internal short-circuit: if they do not, the power pack is in order and the trouble lies elsewhere. Switch off, remake the HT+ connection and pull out all the valves (not the rectifier). Try switching on again. If the fuses blow again, shake the apparatus vigorously to see if any pieces of metal are lodged among the wiring. Suspect composition resistors comprising HT potential dividers—perhaps one which normally constitutes most of the divider resistance has gone 'low'. Any divider resistor which is discoloured, or has cracked paint, has been subjected to overheating and may be the offender.

If the dividers seem to be in order, the trouble is probably a valve taking a greatly excessive current, most likely a power valve. Replace the valves one by one until the culprit is found. With the offending valve pulled out again, check the associated circuitry, looking particularly for excessive screen potentials or insufficient bias.

(2) *Partial failures*

The method here is to check the operation of the unit stage by stage, starting at one 'end' and working towards the other. With apparatus containing pulse circuits it is best to monitor the waveforms with an oscilloscope, beginning at the 'front': thus if the defective piece of gear is a stimulator, begin by checking the output of the relaxation oscillator, then verify that the delay circuits are being triggered properly by the oscillator, then that the shock generator is being triggered by the delay circuits, and so on towards the output terminal until the defective stage is found. This done, the oscilloscope may be exchanged for a voltmeter, and the relevant valve potentials checked. Potential checking is made much easier if one knows what the reading ought to be, and to this end it is a valuable, if monotonous, practice to measure and note down in a special log book the potential readings to earth from all pins of all valves of all newly completed apparatus, *together with the ohms-per-volt of the meter used*. In general an instrument of at least 10,000 Ω/V is advisable. Any dial settings which affect these potentials should, of course, also be noted.

If in the defective stage one of the voltages is clearly much other than it should be, the faulty component—anode load, bias resistor, etc.—should be deducible. If all seems to be correct, look for open-circuited coupling components, e.g. capacitors or germanium diodes.

If the faulty piece of apparatus is an amplifier it is often better to work in the opposite direction, i.e. from the output back towards the input. This is because an oscilloscope is already present in the form of the cathode ray tube which is fed by most biological amplifiers. Using a sine-wave oscillator provided with a calibrated attenuator, inject, via a large capacitor (say 4 μF paper), a signal direct on to one of the *Y* deflector plates of the tube. If the trace produced is of steady amplitude and brightness, the tube supply circuits are in order and the tube can be calibrated in terms of mm/V. Armed

with this information, transfer the oscillator output to the output stage grid (or one of the grids, if a differential amplifier), still using the 4 μF capacitor, and check the output stage gain, distortion, signal handling ability, etc. Continue in like manner towards the amplifier input until the defective stage is found. Once again, the work will be easier if the stage gains to be expected are kept in a log book. Remember that if the amplifier gain control works by variable negative feedback, the gain figures at low-gain settings will be rather independent of the condition of the valves. Only the high-gain values are of much significance.

Having established the defective stage, use the voltmeter to check the valve potentials as before.

(3) *Intermittent faults*

The remedying of intermittent faults is among the most exasperating of the tasks which befall anyone concerned with the maintenance of electronic equipment. With the fault 'on', one begins one's logical chain of experiments and deductions, eventually making some alteration which seems to put matters right, only to find later that the same trouble reappears, and to be forced to the regretful conclusion that one has been on quite the wrong track.

Intermittent faults are usually caused by components on the verge of breakdown, which alternate between normal and abnormal working. If the apparatus can be left on for long enough the breakdown will probably become final and permanent in the end, when the problem can be attacked in the normal manner. Unfortunately it is not always convenient merely to sit down and wait.

It may be possible to accelerate the final demise, or at least to cause the fault condition to appear at will, by a little gentle prodding and tapping at likely components. Alternatively running the apparatus on 10 per cent over-voltage (in gear with unstabilized power packs, for example, by using the 200 V mains transformer tapping on 220 V mains) may be successful, though it is scarcely a practice to be recommended.

There is one type of intermittent fault to which direct-coupled amplifiers are particularly prone, which appears on the cathode ray tube face as a random flickering of the base line between two definite levels. The trouble *may* be caused by a senescent HT battery or faulty power pack, but if this should prove not to be the case it is probably attributable to a defective high-value (e.g. 100 k) wirewound resistor, of which such amplifiers contain a number. The reason seems to be that in course of time the weld between one of the lead-in wires and the resistance element fractures, causing unsatisfactory contact. The equipment needed to find the faulty resistor is quite simple. If the amplifier is single-sided, take a good quality 4 μF paper capacitor and earth one terminal. Connect the other to a yard or two of flex, ending in a crocodile clip. If the amplifier is double-sided, two such capacitors are needed. Any part of the amplifier circuit to which a clip is attached now becomes 'frozen' in potential, so far as rapid changes are concerned.

The procedure is, with the amplifier input short circuited, to work steadily through from the output end towards the input, 'freezing' the signal path, until the stage containing the faulty resistor is found. Begin at

EQUIPMENT FAULTS

the output stage anode (or anodes), then transfer to the output stage grid, then to the penultimate stage anode, and so on. At first the application of the clips will suppress the fluctuations of the base line, but at length a point will be reached at which hooking them on makes no difference. Thus if this point is the second stage grids, suspect the second stage anode loads or cathode resistors. To find out which component is the offender, disconnect the capacitors and connect one of them *across* each suspect resistor in turn. Resistors which, when so shunted, have no effect on the fluctuations may be regarded as in order; those which, when shunted, suppress the fluctuations, or at least round them off, should be replaced by good components in turn until the faulty resistor is found. The culprit so detected may seem to behave quite normally if tested on an ohmmeter, but if replacement proves it in fact to have been responsible, it should be at once destroyed, preferably with a hammer*. It is a bad practice to allow dubious components to lie about where they may be 'rescued' and used by someone else.

The same 'freezing' process may be used to detect the point of entry into the amplifier of 50 or 100 cycle interference from within, e.g. because an electrolytic decoupling or smoothing capacitor has become 'low' or open-circuited.

(4) *Illusory faults*

The purpose of this section is really to utter a warning: not to assault the apparatus with screwdriver and soldering iron until it is clearly established that a fault condition really exists. So often, hours are wasted whilst a 'fault' is eventually found to be an accumulator in need of charging, or a HT battery plug in the wrong tapping, or even not plugged in at all. As further examples of the kind of thing which can baffle one for a time, a stimulator whose delay period is accidentally set to longer than its repetition time will seem to behave in an irrational manner. D.c. amplifiers may refuse to balance simply because of an excessive standing potential developed by the electrode system. Double-sided amplifiers are often used with a single electrode, the other input being earthed. In this case remember that the standing potential developed by the preparation-earth connection must be taken into consideration too. Enormous and inexplicable base-line shifts may be produced in direct-coupled apparatus in sympathy with the movements of the experimenter, because his shadow is falling across his pre-amplifier valves. Some valves, excellent in other respects, are extremely photoelectric, and should be shut up in light-tight boxes. Moving-iron meters can give utterly meaningless readings because they are being used in the neighbourhood of a magnetic field. The list could be extended indefinitely, but to end this somewhat depressing chapter, I retail this story of Professor Matthews'.

A piece of apparatus was set up, away from home ground, for a demonstration. It included a pre- and a main amplifier; the former was stood upon the top of the bench, and the latter on the floor. The cable joining them was passed through a hole, which happened to have been drilled in the bench-top on some previous occasion. Upon trying out the apparatus, the

* Psychologically satisfying.

FAULTS

rise time at the output for a step-function at the input was found to be a few orders too long. Attention was at once directed to the amplifiers, but no explanation could be found. At length, someone spotted a drawer under the bench and below the hole, and inside lay a brand-new 100 yard coil of cable. One end went up through the hole to the pre-amplifier, the other over the back of the drawer to the main amplifier. The technician had not liked to cut it!

DESIGN PROCEDURE

The procedure to be described in this chapter is only one of many possible. It is the way the author works, but he anticipates that readers will soon develop their own methods of attack. The steps which lead from the original conception of a new piece of apparatus to its final realization in the metal are illustrated by the use of an example, the production of a complete electrophysiological unit for students' use.

We begin with a specification, which may run something like this:

Stimulator—two repetition rates, 25/sec and 1/sec; two shocks to be provided, each with a delay variable up to 30 msec from the initiation of the C.R.T. time-base sweep. Exact shock waveforms immaterial, but shock-strength adjustable up to a maximum equivalent to a rectangular pulse of 30 V amplitude into 10 k Ω and lasting $\frac{1}{2}$ msec.

Time base—triggered by stimulator, providing trace lengths adjustable from 10 to 80 msec. Switchable 10 msec time-marking pips to be provided so that the trace length may be set up to an exact multiple of 10 msec within the range provided. Pips derived from the 50 cycle mains would be sufficiently accurate.

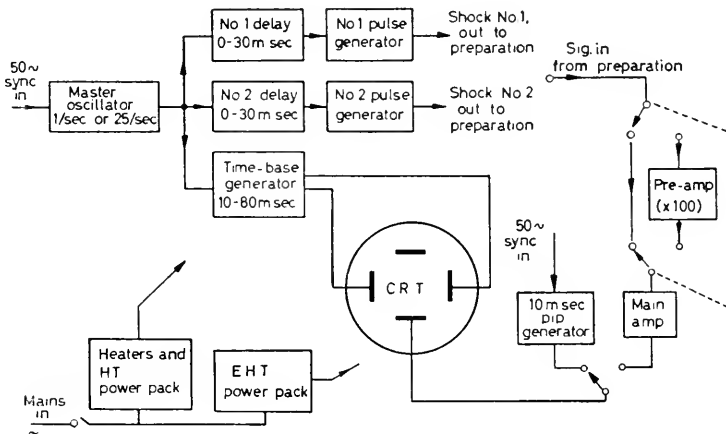


Figure 44.1

Amplifier—A direct-coupled amplifier of maximum gain such that the cathode ray tube spot is deflected to top or bottom of the screen by ± 25 mV: this is for intracellular work. In addition, a pre-amplifier, a.c. coupled, of short time constant, gain 100 times, suitable for the detection of weak action potentials.

Presentation—6 in. diameter flat-faced cathode ray tube, medium persistence (i.e. green) phosphor.

Power supplies—As requisite.

DESIGN PROCEDURE

From the specification, a provisional block diagram of the proposed apparatus follows at once (*Figure 44.1*). The only point to notice at this stage is that if time marks are required for direct visual inspection, it is essential that they appear similarly in successive traces; that is, the train of pips must not march to right or left, otherwise setting up the trace length becomes extremely difficult. Thus, if the time marks are to be derived from the mains, the master oscillator must be synchronized with the mains also.

We now consider the contents of each block in more concrete terms. Initially we propose well-established and trustworthy circuitry, perhaps as follows:

Master oscillator	thyatron relaxation oscillator	1 valve
Time base	sanatron (2 pentodes) and paraphase amplifying valve to derive push-pull output, say another pentode	3 valves
Delay circuits	double triode flip-flops	2 valves
Pulse generator	If the shape of the output pulse is not critical, a very simple possibility here which gives a stimulus moderately free from earth is an electronic version of the traditional induction coil; i.e. a power valve is used to interrupt the current in the primary of a loose-coupled mutual inductance, to the secondary of which the preparation is connected. So for two channels	2 valves
Main amplifier	It would be unwise not to use double-sided amplification here. The cathode ray tube which it is proposed to use has a deflection sensitivity of $1,100/VA_3$ mm/V, so with a final anode potential of 2,000 V, the deflecting voltage required to sweep the full 6 in. is about 300. The gain required is therefore 6,000. A good pentode can give a gain of 150, but not under low-noise conditions, when a figure of 30 is nearer the mark. In addition, since the amplifier is direct-coupled, gain may be lost in the interstage couplings. 6,000 is an awkward amount. It is just too much for 2 stages, but rather too easy for 3. It will be possible to use considerable negative feedback. A reasonable division of labour would be to secure (including couplings) a gain of 15 from the 1st, 20 from the 2nd and 30 from the 3rd.	Total 6 valves
Pre-amplifier	This too will be double-sided. A gain of 100 is beyond the capability of one low-noise stage, but may easily be achieved by two.	Total 4 valves
		Grand Total 18 valves + rectifiers if used

At this point it is possible to produce a first estimate of material cost. As a rough guide, which is about right at the time of writing, the author allows £2 per valve, with its associated circuitry, + £5 for sheet metalwork + £5 for each power pack + cost of cathode ray tube (£20). At present, then, it looks as if the unit will cost about £70.

The next stage is to see if there are any catches in the scheme as so far proposed, and if so to modify it as necessary. For example, the thyatron

master oscillator; the thyatron HT should not be applied until its heater has warmed up. It seems a pity to use a thermal delay switch just for one valve. We might use an indirectly heated thermionic rectifier, which also takes some time to warm up. Or we might provide a 'Thyatron HT switch' to be operated by the user, say, 30 seconds after closing the main switch. Perhaps it would be safer not to use a thyatron at all; what about a blocking oscillator?

Spotting latent difficulties such as these can save trouble later on, but it seems that, however hard one tries, a few always manage to get by unnoticed.

The next phase is to ponder upon the work so far, because it is at this juncture that ideas come which can lead to important simplifications, and may even be the germ of an original piece of circuitry. For example, in the time-base generator, if we applied the trigger signal to the sanatron control valve at the control grid, instead of at the suppressor, we could use a triode instead of a pentode: and if a triode were to prove a satisfactory alternative to a pentode for the paraphase amplifier, then we could use one double triode instead of two pentodes. Net profit—one valve saved. Again, in the stimulator, suppose that instead of having—in each channel—a double triode for the delay flip-flop and a power pentode for the output, we use the screen and control grid of the power pentode *as* one half of the flip-flop. Then one triode-pentode valve would perform the duty both of delay stage and output stage.

Having arrived at a few notions of this kind one is naturally anxious to try them out, and the time is now ripe to begin practical work. For circuit development one needs a good stock of reliable old components—perhaps stripped from obsolete equipment; it is a pity to use new parts here. In addition, certain items of electronic test equipment are required. In the author's opinion the following are essential:

- (1) A good measuring-oscilloscope (i.e., one with the *Y* shift calibrated in volts and the *X* shift in time) provided with direct-coupled amplifiers.
- (2) An audio-frequency oscillator, range about 10–100,000 c/s, with calibrated attenuator, and output switchable to either sine wave or square wave.
- (3) Two multi-range universal meters (i.e., combined voltmeter, ammeter, and ohmmeter, a.c. and d.c.) of which at least one should be of high sensitivity—say 20,000 ohms/volt.

Since the design of power packs for apparatus is best left till last, when the output and degree of stability required are known, a wide-range variable HT supply having at least 3 separate outputs will be found invaluable. Failing this, keep a stock of half a dozen HT batteries of the type well endowed with tappings. Specimens discarded as too old for reliable amplifier operation are commonly quite satisfactory for powering developmental pulse circuits. A 6V car battery and charger are also required.

A valuable optional piece of apparatus which can reduce unnecessary wastage of doubtful components, and establish their values accurately when these have to be known, is a resistance-and-capacitance bridge. These may be bought commercially, but they are quite simple affairs and may be home made. M. G. Scroggie (*Radio Laboratory Handbook*. London; Iliffe) has described an excellent instrument which would be very suitable.

DESIGN PROCEDURE

Assuming the necessary components, test gear and tools are to hand, decide on the first part of the apparatus to be developed and jot down the proposed circuit. Insert some likely component values, bearing in mind the design equations and limits given in earlier parts of this book. Happily electronic design is not like solving a simultaneous equation. There is no single correct combination of component values which is the only one which will work: this may be true of parts of a circuit, e.g. filters, where certain values are uniquely determined if others are known, but for the most part considerable latitude is possible. Often ratios are of more importance than absolute values, as in, for example, potential dividers. It follows that electronic design is not a direct process in which every step is inflexibly determined by its predecessors. For most of us there is no question of sitting down with pen, paper, manufacturers' literature and slide rule, doing some calculations, then handing one's technician a completed circuit diagram in the confident knowledge that the circuit will work, exactly as predicted, straight away. A more empirical approach is both easier and safer.

Having arrived at some values for the new apparatus, build it, one stage at a time, and do not move on to the next stage until the first is giving the required performance. The great thing is to *get the stage going*, if only after a fashion. At least one then has a performance of some sort to measure, waveforms to look at, and voltages to check. Thereafter, a proper understanding of the circuit action, coupled with the application of common sense, should suggest what component values to modify in order that the required behaviour is obtained. The only point to watch is that, in one's enthusiasm, some simple rule about component loading is not overlooked. Over-run resistors and chokes generally draw attention to the fact by discoloration and the emission of peculiar smells, or even smoke, but capacitors can break down without warning, and valves may suffer silently for a time, then fail prematurely.

When the stage is working properly, find out how susceptible it is to variations in circuit parameters, particularly supply voltages, so that an estimate of the required power pack stability may be made. If it is a pulse circuit, make sure that it does not 'work, but only just'. For example, if in the electrophysiological unit which is our example the master oscillator ceases to trigger the time base if the former's output pulse amplitude falls by 10 per cent, then the design is clearly unsatisfactory. The effect of production spreads on component values is such that in subsequent attempts to reproduce the apparatus, a weak oscillator is liable to be combined with a refractory time base, and the gear will not work. Even if it does, it may soon cease to as the valves age. The trigger pulse should be made at least twice the minimum required, either by stepping up the oscillator output or by reducing the suppressor bias which holds the sanatron Miller valve cut off.

In building the first working model the mechanical construction should be of 'lash up' form. At the moment we are concerned with the electrical design. Lay-out is unimportant except in so far as haphazard placing of components may lead to instability due to unwanted capacitive or inductive couplings. The object is simply to provide sufficient mechanical support

DESIGN PROCEDURE

for the components to prevent tags making accidental contacts with others, leading almost certainly to short-circuits or non-operation. In addition, fixing everything together gives a measure of portability, enabling the whole job to be put away if more urgent matters intervene. Various methods are possible: everything can be screwed to a piece of wood, or even cardboard, or a large old chassis, already provided with a variety of holes, may be pressed into service. The author solders everything, valveholders, potentiometers and all, to a long tagboard; in this way an experimental stage can be set up in about 5 minutes.

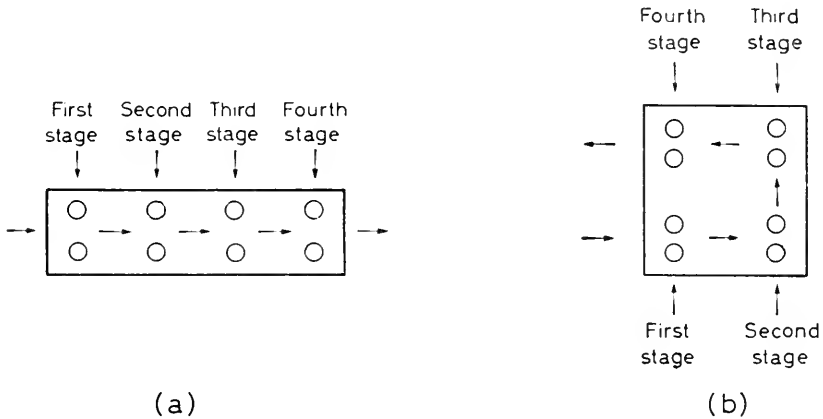


Figure 44.2

Lash ups tend to take up a lot of bench-space, and it would generally be impractical to fashion the entire electrophysiological unit in this way, before building anything in its final form. A reasonable procedure is to develop as a lash up as much as it is eventually proposed to put on one chassis, then do the necessary mechanical design and build it properly before developing the next part. In our example a possible scheme of division might be:

- Chassis 1 Pre-amplifier
- 2 Main amplifier
- 3 Cathode ray tube, with its brightness, focus and astigmatism controls
- 4 Time base and stimulator
- 5 EHT and HT power packs.

Building everything on one huge chassis is, of course, possible, but it may be difficult to get a sufficient separation between, say, the power packs and the pre-amplifier, and the result is anyway likely to be cumbersome.

Methods of mechanical construction have been discussed in Part II. The problem of component layout within the framework of the chosen scheme is one of scale and of pattern. Apparatus laid out on a generous scale is easy to work on and less prone to spurious operation caused by unwanted stray capacitances and inductive couplings, but it is bulky and requires more sheet metalwork. The pattern of layout, also, is bound up with the question of stray couplings. Thus a four-stage differential amplifier should be laid out in plan as shown in *Figure 44.2a* rather than *44.2b*.

DESIGN PROCEDURE

Figure 44.2b may produce a chassis of more convenient shape, but would be extremely liable to oscillate. As a general rule one does not go far wrong by allowing the theoretical circuit diagram to suggest the physical layout. If the panel controls are then arranged to be near the stages they operate on it is usually the case that, since the theoretical circuit is logical, a logical, and therefore conveniently usable, panel layout results automatically. Furthermore the length of wiring runs is minimized. Unfortunately the rule sometimes breaks down with composite valves. For example, what should be done if part of the electrical circuit is as in *Figure 44.3a*? It all depends on circumstances, but *Figure 44.3b* is a possibility.

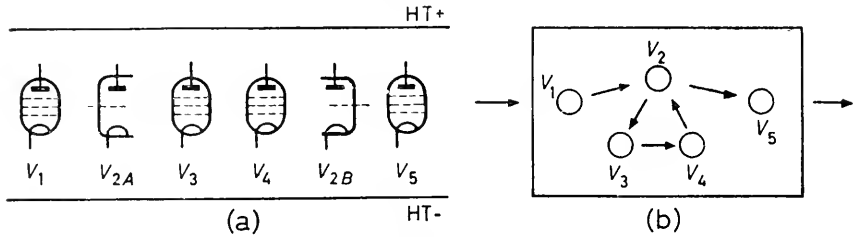


Figure 44.3

When the several chassis have been completed and are bolted together on a rack, it may be found that the various units, though they behave properly alone, do not work properly in concert, due to some kind of interaction, of which the most likely is excessive pick-up by the amplifier circuits of artefacts originating from the time base or stimulator. Setbacks of this kind are to be anticipated and should not be a cause for despondency. Use the 'freezing' technique described in the last chapter to establish how the interference enters the amplifier, whether via the power supplies or by the input circuit, or by some other route. If the former, the output impedance of the power pack must be reduced, or two separate packs used, or if necessary the amplifier may have to be supplied from batteries. If by the input circuit, the methods enumerated in Chapter 43 are available to combat the trouble.

TRANSISTORS

In leaving transistors to the end of this book there is no suggestion that the subject is only of marginal interest to the electrobiologist; rather, this chapter is in the nature of a 'stop press' item. Following the announcement of the invention of the transistor from the Bell Telephone Laboratories¹ of America in 1948 there elapsed an interval of some 5 years for further improvements and the development of the necessary manufacturing techniques, after which transistors with predictable characteristics began to appear upon the market. Once in the hands of the circuit engineers there followed a further period during which the usefulness of the transistor as an active circuit element—and hence as an alternative to the thermionic valve—was investigated. In recent years the results of these endeavours have led to the publication of a veritable spate of transistorized devices. So rapid is the pace of development in this field that it is difficult to write about the subject without one's efforts becoming almost at once out of date.

It would be unwise to predict that the transistor will completely oust the valve in electrophysiological or allied work; indeed, the transistor is at present at a disadvantage compared with the valve in two particulars of great interest to the electrophysiologist. These are: it is difficult to make a transistor amplifier having a high input resistance, and it is also difficult to make a direct-coupled transistor amplifier which is not seriously prone to drift; but as a.c. coupled amplifiers to work from low-resistance preparations, certain pulse generators, high voltage power supplies and voltage stabilizers may all be successfully made with transistors, the subject is clearly deserving of his attention. Advantages of the transistor are small size, absence of microphonicity, excellent power efficiency (no heaters to supply) and ruggedness (transistors can withstand accelerations which would destroy a valve). The question of transistor noise is at present receiving much research effort; it seems probable that future types will surpass valves in this respect but generally speaking those available at present are not as good. The performance at radio frequencies in contemporary transistors is also unsatisfactory, but electrobiology is mainly interested in the frequency band 0–20,000 cycles, so this aspect is of minor importance. At the time of writing transistors are between two and three times as expensive as valves.

Transistors are of two types, 'point-contact' and 'junction'. The point-contact type was announced first, and has certain properties not possessed by the junction transistor which are important in some pulse circuits. Nevertheless, it appears—at least in the field in which we are concerned—that the junction transistor is emerging as by far the more important, and the point-contact type will not therefore be discussed.

In Part II, the use as diodes of junctions between P and N type semi-conductors—germanium or silicon—was mentioned. The triode junction

TRANSISTORS

transistor is a three-terminal device in which two such junctions are arranged back-to-back (*Figure 45.1*). Clearly two dispositions are possible, and these are described as N-P-N, or P-N-P. At present only the P-N-P type is widely used in this country, and we shall concentrate upon this. However, all the remarks which we make about the workings of P-N-P transistors

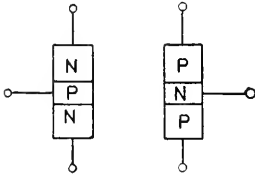


Figure 45.1

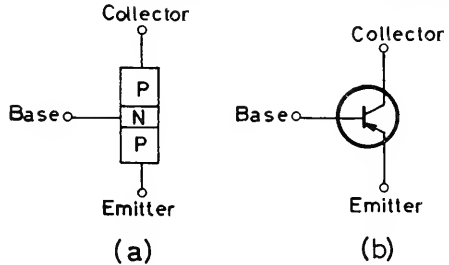


Figure 45.2

apply equally to the N-P-N type, except that in N-P-N circuits the polarity of all supply voltages is reversed, and all currents flow in the opposite direction.

The terminals of a triode* transistor are called the emitter, the base and the collector (*Figure 45.2a*) and may be regarded as corresponding to the cathode, grid and anode of a valve. The circuit symbol for the P-N-P type is shown in *Figure 45.2b*. In the N-P-N type the direction of the emitter arrowhead is reversed. If we regard the transistor as a 'black box' having two input and two output terminals, since the device has actually only three terminals altogether, clearly one of the real terminals must always be common to the input and output circuits, and three configurations are possible:

common base (*Figure 45.3a*)
 common emitter (*Figure 45.3b*)
 and common collector (*Figure 45.3c*)

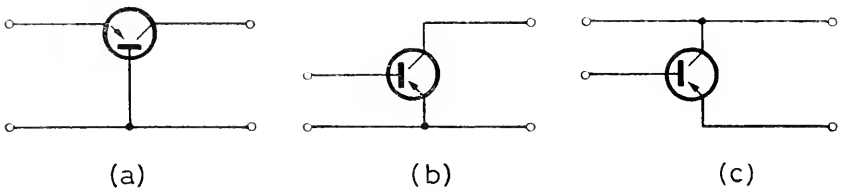


Figure 45.3

Since the common terminal is generally earthed so far as signals are concerned, leaving the remaining two as 'live' input and output connections, these configurations are also called 'earthed base, earthed emitter and earthed collector'. In America the term 'ground' is used for 'earth'. Parenthetically it may be remarked that, in transistor terminology, an amplifier valve is an earthed cathode, and a cathode follower is an earthed anode,

* Tetrode transistors exist, but as yet are a rarity. Henceforward we drop the term 'triode' from the description.

THE EARTHED EMITTER TRANSISTOR

circuit. *Figure 45.3b* is therefore the transistor equivalent of a conventional valve amplifier, and *Figure 45.3c* corresponds to a cathode follower. *Figure 45.3a* is equivalent to a valve circuit which exists but with which we have not dealt—the ‘grounded grid’.

At this point we anticipate the results of explanations which remain to be given and assert that the transistor is, like the valve, an amplifying device but which, unlike the valve, consumes power at the input. From a black-box viewpoint (*Figure 45.4*), we can distinguish the *current gain*, $\delta I_{out}/\delta I_{in}$,

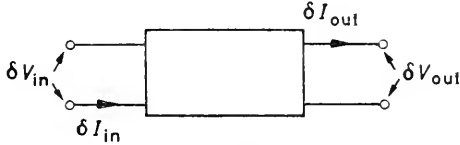


Figure 45.4

the *voltage gain*, $\delta V_{out}/\delta V_{in}$, and the *power gain* $(\delta V_{out} \delta I_{out})/(\delta I_{in} \delta V_{in})$. In the earthed base configuration a transistor gives a voltage gain, but a current gain less than unity. In the earthed collector configuration, a current gain, but a voltage gain less than unity. In the earthed emitter arrangement the device shows both voltage gain and current gain, so that the power gain is maximal: power gains of over 40 dB's are possible. Further, an earthed base transistor has a low input resistance and a high output resistance, higher by about 3 orders, whilst in an earthed collector arrangement the situation is reversed. With the emitter earthed, the input and output resistance are more comparable (though the latter is still at least 10 times greater). This facilitates the coupling of transistors in cascade, since the matching conditions for optimum power transfer are most nearly met.

The earthed emitter configuration is therefore of the greatest interest, and will receive the bulk of our attention. With transistors occupying their present low level of importance in biological work, a full treatment involving all three configurations would produce a chapter of disproportionate length. There is no dearth of textbooks in which common base and common collector configurations are properly discussed; our purpose here is to introduce transistors and to cover sufficient mathematics to enable simple transistor design work to be carried out. Before beginning it must be mentioned that considerable confusion exists at present (September 1957) about the meaning of transistor symbols. A number of groups of transistor engineers have developed their own transistor symbolism and it is as yet too early for standardizing authorities to decide upon, and lay down, the best system. This confusion accounts for much of the difficulty in understanding transistors at the present time.

THE EARTHED EMITTER TRANSISTOR

Suppose we arrange an earthed emitter transistor in the test set-up shown in *Figure 45.5*. We can vary the current flowing out of the base by the variable resistor R_V , seeing what effect is produced on the collector voltage and current, as indicated by the voltmeter and milliammeter. R_1 and R_2

TRANSISTORS

are safety resistors to limit the currents and prevent accidents to the transistor. Then the graph which emerges has the form shown in *Figure 45.6*. Probably the first thing which strikes us about this 'collector characteristic' for the transistor is the resemblance to the anode characteristic of a pentode. Above a certain 'knee' collector voltage, the collector current I_c is determined

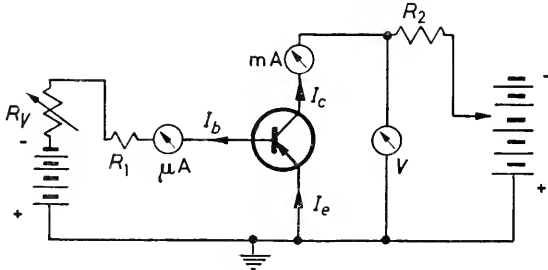


Figure 45.5

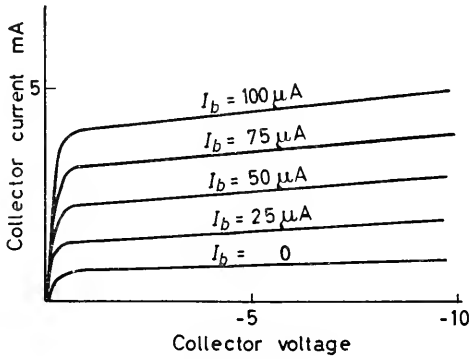


Figure 45.6

almost entirely by the base current I_b and is much larger than it. We can define A , the 'forward current amplification factor' as

$$\left[\frac{\delta I_c}{\delta I_b} \right]_{V_c = \text{const}}$$

It is generally between 20 and 60. This suggests the beginnings of an equivalent circuit for the transistor, *Figure 45.7*. However, above the knee

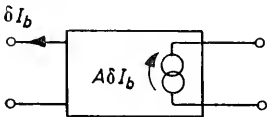


Figure 45.7

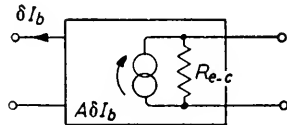


Figure 45.8

the characteristics are not perfectly flat; the reciprocal of their slope represents a resistance R_{c-c} , seen looking into the transistor between emitter and collector, which might in a typical small transistor be 50 kΩ. Our equivalent circuit is modified to *Figure 45.8*.

THE EARTHED EMITTER TRANSISTOR

Now if this were all we had to do to specify the transistor, the subject would be no more difficult than valves; for we have two parameters, R_{cc} which corresponds to r_a of a pentode, and A , which corresponds (but not dimensionally) to g_m . Unfortunately we have to discover what is happening at the input circuit. We keep the investigation symmetrical if we set up

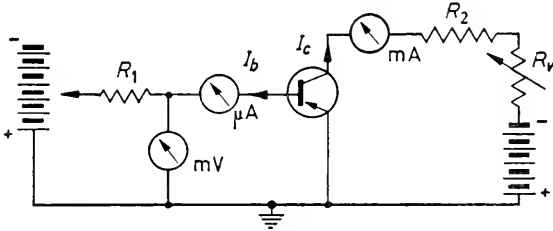


Figure 45.9

Figure 45.9, seeing how the base voltage and current are related to the collector current. On plotting the results we find a characteristic like Figure 45.10. When $I_c = 0$ we have merely the base-emitter junction as a diode biased in the forward direction; voltage across it is low, and the resistance is represented by the reciprocal of the characteristic slope, about $1\text{ k}\Omega$ in the small transistor we are considering. The effect of allowing

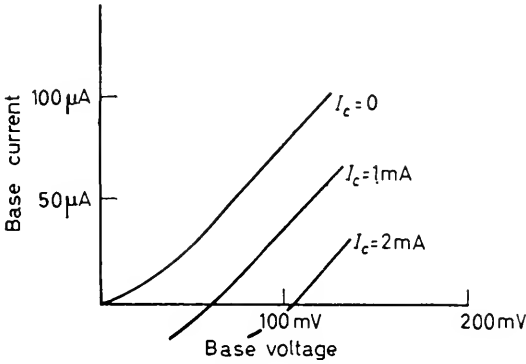


Figure 45.10

collector current to flow is to *reduce* the base current at a given base voltage, and we can define a ‘backward current amplification factor’

$$\left[\frac{\delta I_b}{\delta I_c} \right]_{V_b = \text{const}}$$

and call it B . Its value is less than 1, about 0.04 in the case considered; the transistor does not amplify in both directions at once. Our equivalent circuit now has the appearance of Figure 45.11, and is complete. It is only one of many possible equivalents for a transistor, but has the merit that the four parameters required follow from the characteristic curves.

Another equivalent circuit, simpler in that it contains only one generator, is shown in Figure 45.12. Perhaps its greatest merit lies in the fact that it is

TRANSISTORS

possible to attach some physical significance to the elements involved. Thus R_e is the resistance of the emitter-base junction, which, being a diode biased in the forward direction, is low (40Ω in the transistor considered). R_c is the resistance of the base-collector junction, which is a diode biased in the reverse direction, and is high ($50 \text{ k}\Omega$). R_b is the resistance of the base

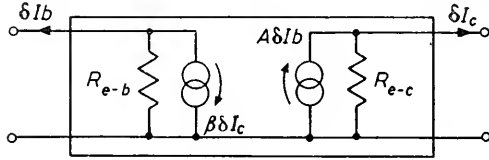


Figure 45.11

material between the external connection and the region between the junctions and is intermediate ($1 \text{ k}\Omega$). To explain the presence of the current generator requires an explanation of transistor action, a perilous voyage into the sea of semi-conductor physics which the author has no wish to make. It is sufficient to take it that the transistor behaves as if there were a current generator in the collector circuit of output β times the base current.

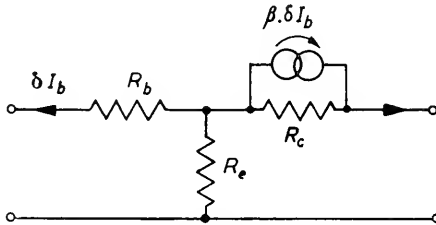


Figure 45.12

β is closely related to A in the previous equivalent circuit and is numerically close to it, between 20 and 60. Other symbols used for the quantity β are α' and α_{b-c} .

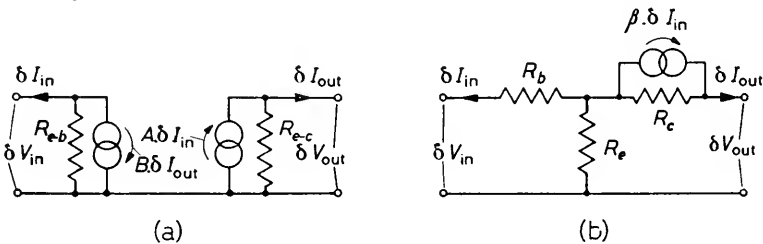


Figure 45.13

It is not difficult to move between the two possible sets of transistor parameters proposed so far. In Figure 45.13a we have

$$\delta V_{in} = R_{e-b}(\delta I_{in} + B\delta I_{out})$$

that is
and

$$\delta V_{in} = R_{e-b}\delta I_{in} + R_{e-b}B\delta I_{out}$$

that is

$$\delta V_{out} = R_{e-c}(\delta I_{out} - A\delta I_{in})$$

$$\delta V_{out} = -R_{e-c}A\delta I_{in} + R_{e-c}\delta I_{out}$$

THE EARTHED EMITTER TRANSISTOR

and in *Figure 45.13b* we have

$$\delta V_{in} = (R_e + R_b)\delta I_{in} + R_e\delta I_{out}$$

and

$$\delta V_{out} = (R_e - \beta R_c)\delta I_{in} + (R_e + R_c)\delta I_{out}$$

Comparing the coefficients of δI_{in} and δI_{out} we see that

$$(R_e + R_b) = R_{e-b}$$

$$R_e = B \cdot R_{e-b}$$

$$(R_e + R_c) = R_{e-c}$$

$$\beta R_c - R_e = AR_{e-c}$$

but with practical transistors it is sufficiently accurate to write

$$(R_e + R_b) = R_{e-b}$$

$$R_c = B \cdot R_{e-b}$$

$$R_c = R_{e-c}$$

$$\beta = A$$

If we provide a transistor with a load resistance R_L and feed it from a

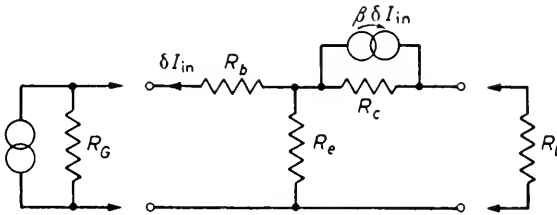


Figure 45.14

signal generator of internal resistance R_G (*Figure 45.14*) then these exact expressions² follow:

The current gain

$$\gamma_i = \frac{\delta I_{out}}{\delta I_{in}} = \frac{\beta R_c - R_e}{R_c + R_e + R_L}$$

The input resistance, R_{in} , that is, the resistance seen looking in at the input terminals with the generator temporarily disconnected but with the load connected, equals $R_b + R_c(1 + \gamma_i)$.

The voltage gain,

$$\gamma_v = \frac{\delta V_{out}}{\delta V_{in}} = -\gamma_i \frac{R_L}{R_{in}}$$

The output resistance, R_{out} , that is, the resistance seen looking back into the output terminals with the load temporarily disconnected but with the generator connected, equals

$$R_c + R_e \left(1 + \frac{\beta R_c - R_e}{R_b + R_e + R_G} \right)$$

TRANSISTORS

With practical transistors it is usually sufficiently accurate to write these as:

$$\begin{aligned}\gamma_i &= \beta \frac{R_c}{R_c + R_L} \\ R_{in} &= R_b + \gamma_i R_e \\ \gamma_v &= -\gamma_i \frac{R_L}{R_{in}} \\ R_{out} &= R_c + \frac{\beta R_c R_c}{R_b + R_c + R_L}\end{aligned}$$

Thus in the transistor we have been considering ($R_e = 40 \Omega$, $R_b = 1 \text{ k}\Omega$, $R_c = 50 \text{ k}\Omega$, $\beta = 30$) the current gain moves from 30 to 0 as we vary the load resistance from 0 to infinity, and the input resistance goes from 1 k Ω to 2.2 k Ω . The output resistance rises from 50 k Ω , when the generator resistance is infinite, to 110 k Ω , when the generator resistance is zero. A typical value of R_L is 5 k Ω . With this value the current gain is $30\{50 \text{ k}/(50 \text{ k} + 5 \text{ k})\} = 27$. The resistance seen looking in at the input of the transistor is 2.1 k Ω . The voltage gain is $27 \times (5 \text{ k}/1.9 \text{ k}) = 71$. The power gain = voltage gain \times current gain = $71 \times 27 = 1,900 = 32.8 \text{ dB's}$.

Simple design procedure

The dependence of the input resistance of a transistor on the load used, and of the output resistance on the generator used, means that in multi-transistor amplifiers the stages cannot be designed in isolation; rather must the circuit be regarded as a whole. We have here another reason for the

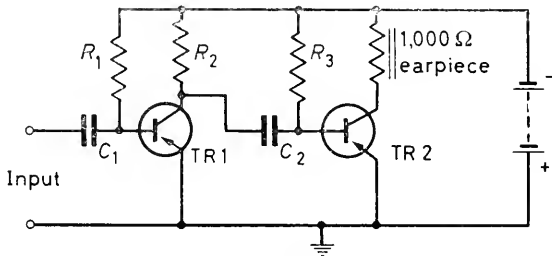


Figure 45.15

greater difficulty in working with transistors as opposed to valves. As an example, consider the simple R-C coupled two-stage audio amplifier in Figure 45.15, in which two transistors in the earthed emitter mode feed an electromagnetic earpiece. We design around two Mullard OC 71 transistors.

Output stage—Starting at the output stage, we begin with the collector characteristic of TR₂ in a manner very reminiscent of valve practice. The OC 71 has a maximum permissible emitter-to-collector voltage of -5 , a maximum permissible collector current of 10 mA, and a maximum permissible collector dissipation of 25 mW. Since the 'knees' of the characteristic curves are at a collector voltage of about -0.25 , we have four figures to define a 'working region' (Figure 45.16). A working point on this diagram for which the makers of the transistor give parameters is

THE EARTHED EMITTER TRANSISTOR

$V_c = -2$, $I_c = 3$ mA. If we were to use this working point the load line for the $1,000 \Omega$ earpiece would be as shown (*LL1*) and cuts the voltage axis at the required battery voltage, in this case 5. If we power our amplifier from dry cells, we cannot have 5 volts, but we can have 4.5 or 6. Let us choose 6. Then we have a new load line *LL2*, and a reasonable working point might be $V_c = -2.5$, $I_c = 3.5$ mA. Then the maximum possible working point excursion is from approximately -0.25 V to -4.75 V

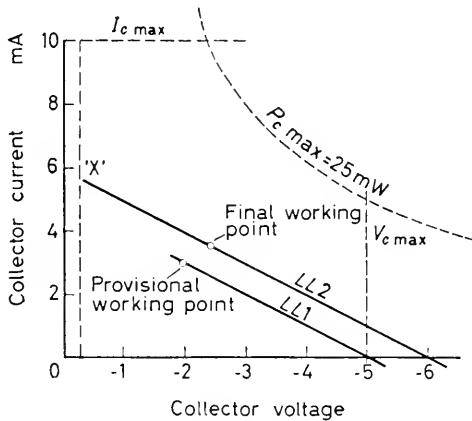


Figure 45.16

(a swing of 4.5 V) and from 1.25 mA to 5.75 mA (a swing of 4.5 mA) so the maximum power output for a sinusoidal signal is

$$\frac{4.5 \text{ V} \times 4.5 \text{ mA}}{8} \approx 2.5 \text{ mW}$$

The power supplied to the output stage is $6 \text{ V} \times 3.5 \text{ mA} = 21 \text{ mW}$, so the efficiency of the stage is $2.5/21 \times 100$ per cent = 12 per cent. Thus far the procedure is exactly as for valves.

In using a working point $V_c = -2.5$, $I_c = 3.5$ mA, our transistor parameters are probably not far from those quoted by the makers for $V_c = -2$, $I_c = 3$ mA, which are

$$\beta = 47$$

$$R_e = 7 \Omega$$

$$R_b = 500 \Omega$$

$$R_c = 13 \text{ k}\Omega$$

To secure a standing collector current of 3 mA we need a base 'bias' current defined by the bias resistor R_3 . The quickest way to find the proper value of R_3 is by experiment with the particular transistor in use. As a pointer to the approximate value required, if we regard β as a measure of I_c/I_b (actually, of course, β is $\delta I_c/\delta I_b$) then to get a collector current of 3 mA we need a base bias current of $3 \text{ mA}/47 = 64 \mu\text{A}$. Since the base and emitter are practically at the same potential, the whole battery voltage appears

TRANSISTORS

across R_3 , which must in consequence be somewhere near $6\text{ V}/64\ \mu\text{A} = 94\ \text{k}\Omega$.

The current gain, γ_i , of the output stage is

$$\beta \frac{R_c}{R_c + R_L} = 47 \cdot \frac{13\ \text{k}}{13\ \text{k} + 1\ \text{k}} \approx 44$$

The input resistance, R_{in} , is $R_b + \gamma_i R_c = 500 + 44 \times 7 \approx 800\ \Omega$

The voltage gain, γ_v , is $\gamma_i \frac{R_L}{R_{in}} = 44 \times \frac{1,000}{800} = 55$

The power gain is $44 \times 55 = 2,400 \approx 34\ \text{dB}$'s

Input stage—The collector of transistor TR_1 sees two kinds of load: so far as the mean working conditions are concerned, the load is the resistor R_2 , but to signals the load is given by R_2 in parallel with the input resistance of TR_2 . A reasonable working point for TR_1 , and one for which the makers of the OC 71 give data, is $V_c = -2\ \text{V}$, $I_c = 1\ \text{mA}$. The transistor not being a completely linear device, the parameters are somewhat different, viz.

$$\begin{aligned}\beta &= 41 \\ R_c &= 18\ \Omega \\ R_b &= 700\ \Omega \\ R_e &= 25\ \text{k}\Omega\end{aligned}$$

For a collector voltage of -2 we have to drop $4\ \text{V}$ across R_2 . If the current is $1\ \text{mA}$, the value of R_2 must be $4\ \text{k}\Omega$. As with the second stage, the quickest way to find R_1 to give the chosen collector current is by experiment.

To signals the load seen by TR_1 is $4\ \text{k}\Omega$ in parallel with $800\ \Omega$, which is $670\ \Omega$. The current gain is therefore

$$41 \times \frac{25\ \text{k}\Omega}{25\ \text{k}\Omega + 670\ \Omega} \approx 40$$

The input resistance = $700\ \Omega + 40 \times 18\ \Omega$
= $1,400\ \Omega$

The voltage gain is $40 \times \frac{670}{1,400} \approx 19$

The power gain is $40 \times 19 = 760 = 29\ \text{dB}$

Coupling capacitor—To compute the required coupling capacitor we refer to *Figure 45.17a*; looking back from the capacitor into TR_1 we see R_{out} , the output resistance of the transistor in parallel with R_2 , the static load. Looking forward from the capacitor into TR_2 we see R_3 , the bias resistor for TR_2 in parallel with its input resistance R_{in} . The time constant of the coupling is given by CR_T , where $R_T = (R_{out} \text{ in parallel with } R_2) + (R_3 \text{ in parallel with } R_{in})$. The value of R_{out} depends upon the resistance of the signal source feeding TR_1 , but it cannot be lower than R_c for the

THE EARTHED EMITTER TRANSISTOR

transistor, which is 25 k. R_3 is much too high to have any appreciable effect on the parallel resistance of R_3 and R_{in} , and may be ignored.

$$\therefore R_T = \frac{4 \text{ k} \times 25 \text{ k}}{(4 \text{ k} + 25 \text{ k})} + 800 \Omega = 4.3 \text{ k}\Omega$$

For the amplifier to operate down to a 3 dB point at ω_c radians per second, we have $1/(CR_T) = \omega_c$, so for our amplifier to go down to 50 cycles, $\omega_c = 2\pi \times 50 = 310$, and

$$C = \frac{1}{\omega_c R_T} = \frac{10^6}{310 \times 4.3 \times 10^3} \mu\text{F} \\ = 0.75 \mu\text{F}$$

Notice the large value required compared with in valve circuits. Fortunately the working voltages involved are low, and physically small electrolytic

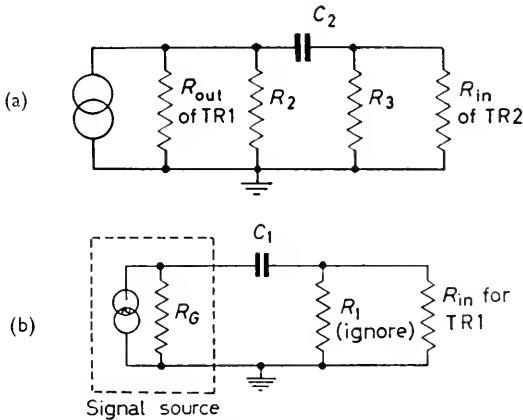


Figure 45.17

capacitors may be used in this position. The value of the input capacitor C_1 may be worked out in similar fashion, from the diagram in Figure 45.17b.

Transformer coupling—The total power gain for the amplifier is $29 + 34 = 63$ dB's. The low figure obtained from the first stage is attributable to the low value of load presented to the first stage by the second. Matters can be greatly improved by using a transformer to match the stages together, though it should be observed that three R-C coupled transistor stages may have the same performance as, but be lighter, smaller and cheaper than, two transformer-coupled stages.

The output resistance for TR_1 is given by

$$R_c + \frac{\beta R_e R_c}{R_b + R_c + R_G}$$

and lies between R_c (when the source resistance R_G is ∞)

and $R_c + \frac{\beta R_e R_c}{R_b + R_c}$ (when $R_G = 0$)

TRANSISTORS

In our particular case it lies between 25 and 50 k. To keep the amplifier general in purpose we do not want to tie down R_o too tightly. A reasonable procedure is to take the output resistance as the geometric mean of 25 and 50 k, ≈ 34 k, and design a transformer so that the second stage presents a matching load resistance of 33 k to the first.

The circuit might have the appearance of *Figure 45.18*. R_d and C_d are decoupling components and R_d drops the battery voltage from 6 to 2 for

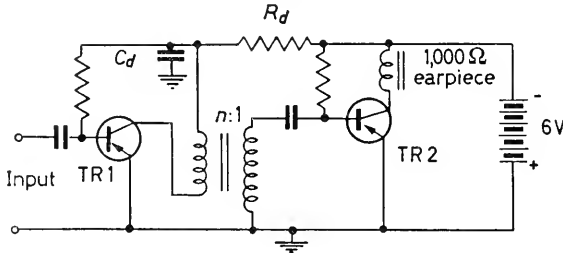


Figure 45.18

TR₁. It is assumed that there is negligible voltage drop across the primary resistance of the transformer. Then the turns-ratio required for the transformer is

$$\begin{aligned} & \left(\frac{R_{\text{out for TR}_1}}{R_{\text{in for TR}_2}} \right)^{1/2} \\ & \left(\frac{34 \text{ k}}{800 \Omega} \right)^{1/2} \\ & = 6.5 \text{ to } 1 \end{aligned}$$

We can now re-assess the performance of TR₁. The current gain, γ_i , is

$$\beta \cdot \frac{R_C}{R_C + R_L} = 41 \times \frac{33 \text{ k}}{33 \text{ k} + 33 \text{ k}} = 21$$

The input resistance is $R_b + \gamma_i R_e$

$$= 700 \Omega + (21 \times 18 \Omega) = 1,100 \Omega$$

The voltage gain, γ_v , is $\gamma_i \times \frac{R_L}{R_{\text{in}}}$

$$\begin{aligned} & = 21 \times \frac{33 \text{ k}}{1.1 \text{ k}} \\ & = 630 \end{aligned}$$

The power gain is $\gamma_i \gamma_v = 21 \times 630 = 13,000$

$$= 41 \text{ dB's}$$

The total gain of the amplifier is now $41 + 34 = 75$ dB's. We have secured a 12 dB improvement merely by paying attention to the matching conditions.

Temperature effects—The amplifiers in *Figures 45.15* and *45.18*, whilst serving to illustrate the design procedure so far as gains, and input and

THE EARTHED EMITTER TRANSISTOR

output resistances are concerned, might not be satisfactory in practice; the performance would be highly dependent upon the ambient temperature. If a transistor be connected as shown in *Figure 45.19*, a small base-to-collector leakage current flows which is attributable to the presence of a few unbound electrons within the lattice of the semi-conductor. This is simply the same as saying that the backward resistance of a semi-conductor diode is not infinite. The leakage current is called $I_{c(o)}$, and is of the order of

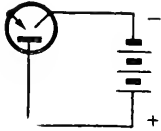


Figure 45.19

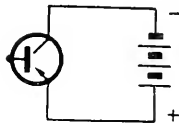


Figure 45.20

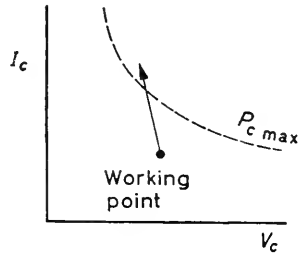


Figure 45.21

10 μA in a small transistor with a typical base-collector voltage at room temperature.

If now the transistor be connected in the grounded-emitter mode (*Figure 45.20*) with the base open-circuited, the leakage current is found to be about 30 times greater. The leakage current for the base-collector junction has to come via the emitter-base junction. So far as the transistor is concerned, it is as if the emitter-base current were due to a base current being drawn off by an external input circuit; transistor action occurs and an emitter-collector current β times the emitter-base current flows. The total collector current is $(1 + \beta)I_{c(o)}$, and is evident in *Figure 45.6* as the collector current which flows when $I_b = 0$. Unfortunately $I_{c(o)}$ is highly dependent upon the temperature of the base-to-collector junction. In the small transistor considered $I_{c(o)}$ might rise to 50 μA at a junction temperature of 50°C.

Consider the effect of this on the output stage of the amplifier in *Figure 45.15*. At a junction temperature of 50°C an emitter-collector leakage current of 50 μA $(1 + 47) = 2.4 \text{ mA}$ flows, producing an additional voltage drop across the 1 k Ω load of 2.4 V, and moving the working point to *X* in *Figure 45.16*. On superposing input signal current on the standing base bias current, the working point can move along the load line downwards and to the right, but not upwards and to the left. The output waveforms are 'clipped' and the transistor is said to be 'bottomed'.

In transformer-coupled stages the effect is much more serious. Here the load resistance, so far as the steady operating conditions are concerned, is very low, being merely the primary resistance of the transformer. The effect of increase of temperature is to move the working point towards the hyperbola of maximum collector power dissipation. Because the collector voltage is approximately maintained, the rising leakage current across the base-collector junction generates heat which further raises the junction temperature. A cumulative action—'thermal runaway'—occurs which eventually destroys the transistor (*Figure 45.21*).

Of the possible circuits to compensate for these temperature effects, the most widespread is shown in *Figure 45.22*. A stabilizing resistor R_s is connected in the emitter circuit, and the base bias is derived from a potential divider R_1 - R_2 across the battery. If R_1 and R_2 are chosen so that the current up them is much larger than the base current, then the mean potential of the base below earth is merely $\{R_2/(R_1 + R_2)\}V$. Since the base-emitter voltage with transistors is small, about 100 mV, the expression

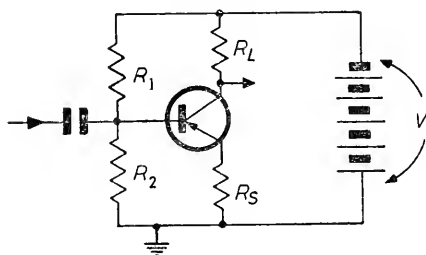


Figure 45.22

$\{R_2/(R_1 + R_2)\}V$ also gives with sufficient accuracy the voltage across R_s , defining the emitter current at

$$\frac{\frac{R_2}{R_1 + R_2} V}{R_s}$$

Of this, a fraction given very roughly by $1/\beta$ flows out of the base, and the remainder out of the collector. The total collector current is composed of the true transistor collector current βI_b , plus the leakage current $(\beta + 1)I_{c(o)}$. Any tendency for the leakage current to increase produces an increase in the potential difference across R_s and a reduction in the emitter-base voltage. Because the emitter-base junction is a diode biased in the forward direction, the voltage-current relation is exponential and a small proportional reduction in voltage produces a much larger proportional reduction in base current. Thus an increase in $(\beta + 1)I_{c(o)}$ is partially offset by a reduction in βI_b .

At first sight it appears that the compensation breaks down when βI_b is reduced to zero, and that therefore the collector current can never be less than $(\beta + 1)I_{c(o)}$. In fact this is not so, because of the peculiar shape of the base-emitter characteristic for transistors in the earthed-emitter mode. It is clear from *Figure 45.10* that the direction of the base current reverses before the emitter-base voltage has fallen to zero. Thus as $(\beta + 1)I_{c(o)}$ increases, a point is reached at which base current begins to flow into, instead of out of, the transistor. βI_b changes sign and the value of collector current is maintained.

An exact analysis of temperature compensation circuits is complicated and cannot be undertaken here. Oakes³ and Stuart-Monteith⁴ may be consulted. In brief, the greater R_s in comparison with the other circuit resistance, the better the stabilization. Compensation for the output stage of the amplifier in *Figure 45.15* may be carried out along approximate lines as follows:

THE EARTHED EMITTER TRANSISTOR

Suppose we increase the battery voltage to 7.5, and lose the extra 1.5 V across R_s . We wish to keep the same working point, $V_c = -2.5$ volts, $I_c = 3$ mA. Then for an emitter current (approximately the same as the collector current) of 3 mA, we need $R_s = 1.5 \text{ V}/3 \text{ mA} = 500 \Omega$. The emitter potential will then be 1.5 V negative to earth, and the base about the same. Therefore $7.5 \cdot \{R_2/(R_1 + R_2)\}$ must equal 1.5. Suitable values would be $R_2 = 10 \text{ k}$, $R_1 = 40 \text{ k}$, giving the circuit of *Figure 45.23a*.

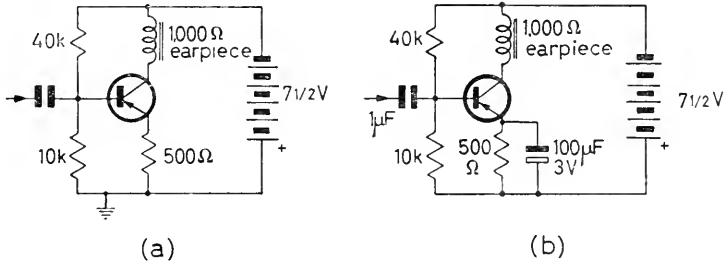


Figure 45.23

There remains one thing to do. So far as signals are concerned, R_s is part of the load in a manner analogous to the cathode resistor in a concertina phase-splitter with valves, and introduces negative feedback. Where such feedback is undesirable it may be removed by shunting R_s with a low impedance to alternating currents; an electrolytic capacitor of the order of $100 \mu\text{F}$ is suitable, to give *Figure 45.23b* as the complete circuit for the stage. The base potential-divider slightly reduces the input resistance of the stage, since the 10 and the 40 k resistors are both effectively in parallel with it; a rather bigger coupling capacitor is therefore called for. Some input signal power is lost in R_1 and R_2 ; their values should therefore not be chosen too low, but equally they must not be too high or the stabilization can be shown to be adversely affected. The values arbitrarily allotted are typical.

Other transistor parameters and equivalent circuits

We now enumerate some of the other ways of considering transistors which are extant, showing how to equate the alternative approaches to that used in this chapter.

The current generator in *Figure 45.12* is sometimes replaced by a voltage generator of output $\delta I_{in} R_m$ (*Figure 45.24*). R_m is a 'mutual resistance';

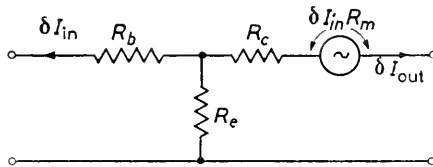


Figure 45.24

that is, the transistor is regarded as introducing so many volts in the collector circuit per amp in the base circuit (*cf.* mutual conductance, so many amps in the anode circuit per volt at the grid, with valves). The necessary conversion factor is just $R_m = \beta R_c$.

TRANSISTORS

It is perhaps unfortunate that although the earthed emitter circuit is probably the most important, many manufacturers describe their transistor parameters in terms of the (possibly more fundamental) *earthed-base* circuit (Figure 45.25a). Fortunately the conversion factors to get from Figure 45.24 to Figure 45.25a are quite simple:

$$\begin{aligned} r_b &= R_b \\ r_e &= R_e \\ r_c &\approx \beta R_c \\ r_m &= R_m \end{aligned}$$

Alternatively, the common-base configuration may be used with a current generator (Figure 45.25b) where $\alpha r_c = r_m$.

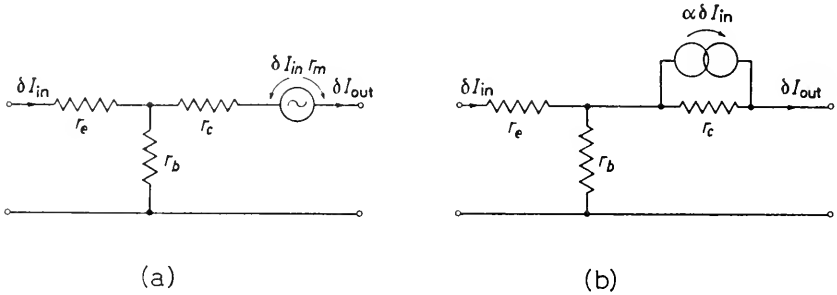


Figure 45.25

Thus, transistors may be described in terms of r_b , r_c , r_e and r_m or r_b , r_c , r_e and α . In one work of reference known to the writer, r_b , r_e , r_c and β are used, though this seems a little inelegant, since the first three belong to the earthed base configuration, and the last to the earthed emitter. The exact relationship between β and α is $\beta = \alpha/(1 - \alpha)$; α is slightly less than unity.

Alternative schemes of description which follow regard the transistor as a black box (Figure 45.26), and different sets of parameters are required

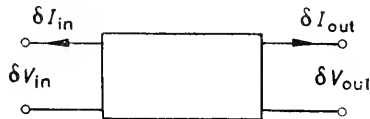


Figure 45.26

for each of the three transistor configurations. It is important not to confuse the sets. One method used for labelling them is to leave earthed base parameters unprimed, to give earthed emitter parameters one prime, and earthed collector parameters two primes.

The four-resistance system—This describes the transistor in terms of two resistances and two mutual resistances or ‘trans-resistances’:

$$\begin{aligned} r_{11} &= \left[\frac{\delta V_{in}}{\delta I_{in}} \right]_{\delta I_{out} = 0} \\ r_{21} &= \left[\frac{\delta V_{out}}{\delta I_{in}} \right]_{\delta I_{out} = 0} \end{aligned}$$

THE EARTHED EMITTER TRANSISTOR

$$r_{12} = \left[\frac{\delta V_{in}}{\delta I_{out}} \right] \delta I_{in} = 0$$

$$r_{22} = \left[\frac{\delta V_{out}}{\delta I_{out}} \right] \delta I_{in} = 0$$

The hybrid parameter system—Describes the transistor in terms of a resistance, a conductance, and two ratios, as follows:

$$h_{11} = \left[\frac{\delta V_{in}}{\delta I_{in}} \right] \delta V_{out} = 0$$

$$h_{22} = \left[\frac{\delta I_{out}}{\delta V_{out}} \right] \delta I_{in} = 0$$

$$h_{21} = \left[\frac{\delta I_{out}}{\delta I_{in}} \right] \delta V_{out} = 0$$

$$h_{12} = \left[\frac{\delta V_{in}}{\delta V_{out}} \right] \delta I_{in} = 0$$

Hybrid parameters are sometimes used with a different set of subscripts, i.e.

$$h_i = h_{11}$$

$$h_o = h_{22}$$

$$h_f = h_{21}$$

$$h_r = h_{12}$$

The Mullard parameter system—By giving 5 parameters in all, so that one is theoretically redundant, this system is in the author's opinion the most useful of the 'black box' descriptions. The maximum current gain and the limits of output and input resistance are at once evident. The parameters are:

$$\alpha = \left[\frac{\delta I_{out}}{\delta I_{in}} \right] \delta V_{out} = 0$$

$$r_{11} = \left[\frac{\delta V_{in}}{\delta I_{in}} \right] \delta I_{out} = 0$$

$$r_{in} = \left[\frac{\delta V_{in}}{\delta I_{in}} \right] \delta V_{out} = 0$$

$$r_{22} = \left[\frac{\delta V_{out}}{\delta I_{out}} \right] \delta I_{in} = 0$$

$$r_{out} = \left[\frac{\delta V_{out}}{\delta I_{out}} \right] \delta V_{in} = 0$$

Thus, in the earthed emitter mode, the Mullard parameter α' is the quantity we have called β . r_{11} is the input resistance when the load is infinite, and is our $R_e + R_b$. r_{in} is the input resistance when the load is zero, and is our $R_b + \gamma_i R_e$ where $\gamma_i = \beta$. Similarly, r_{22}' is the output resistance when the

TRANSISTORS

generator resistance is infinite, and is our R_G ; and r_{out}' is the output resistance when the generator resistance is zero, and is our

$$R_c + \frac{\beta R_c R_c}{R_b + R_c + R_G} \text{ where } R_G = 0$$

Working entirely in terms of Mullard parameters, the resistance seen looking in at the input of a transistor in earthed emitter feeding a load R_L is:

$$r_{11}' \cdot \frac{R_L + r_{out}'}{R_L + r_{22}'}$$

The resistance looking into the output of a transistor fed from a signal generator of resistance R_G is

$$r_{22}' \cdot \frac{R_G + r_{in}'}{R_G + r_{11}'}$$

The current gain is

$$\alpha' \cdot \frac{r_{22}'}{R_L + r_{22}'}$$

and the voltage gain is

$$\frac{\alpha' r_{22}' R_L}{r_{11}' (R_L + r_{out}')}$$

The power gain is, of course, the product of the voltage and current gains.

EARTHED COLLECTOR CONFIGURATION

We now mention briefly the performance of the earthed collector stage, an a.c. coupled version of which is shown in *Figure 45.27* and which, it will

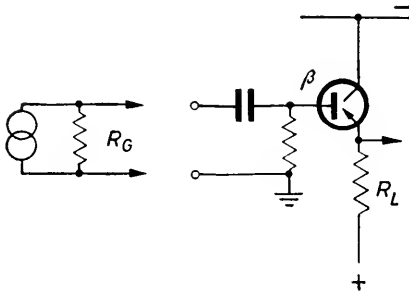


Figure 45.27

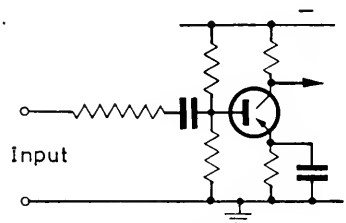


Figure 45.28

be remembered, corresponds to the cathode follower in valve practice. The fact that the input resistance can be arranged to be high (100 k Ω) makes it appear at first sight that this configuration is the most suitable for the input stage of a transistorized biological amplifier; there would be reasonable efficiency of signal transfer from the lower resistance type of electrode system (e.g. nerve lying over silver-wire hooks, large surface electrodes, etc.). Further, the low output resistance makes the configuration appropriate for driving low resistance loads, such as meters.

EARTHED COLLECTOR CONFIGURATION

Approximate equations for the earthed collector stage are very simple. The current gain is roughly β , but the voltage gain is less than unity. The input resistance is merely βR_L and the output resistance R_G/β . Thus if R_L is 10 k and $\beta = 30$, the input resistance is about 300 k Ω . If the earthed collector transistor is being used as an output stage, R_G is the output resistance of the penultimate stage as a whole and is substantially the resistance of its collector load, perhaps 3 k Ω . The output resistance of the earthed collector stage is then 3 k $\Omega/30 = 100 \Omega$.

In point of fact, high input resistance is not of first importance with grounded-collector stages. The power gain is so low that, as Jones and Hilbourne⁵ point out, the performance is no better than that of an earthed emitter stage with the input resistance raised by the simple expedient of connecting a resistance in series with it (*Figure 45.28*). The low output impedance therefore remains its most valuable feature.

Gain control

In potentiometric gain control between valve stages V_1 and V_2 (*Figure 45.29*), the potentiometer is working between a generator of resistance,

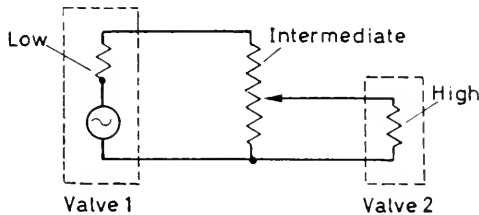


Figure 45.29

the output resistance of V_1 , perhaps 50 k Ω , and a load which is much higher, perhaps many hundreds of megohms. The potentiometer's own value is intermediate, say 2 M Ω , and we saw in Chapter 2 that such an arrangement produces a satisfactory control characteristic. In transistor

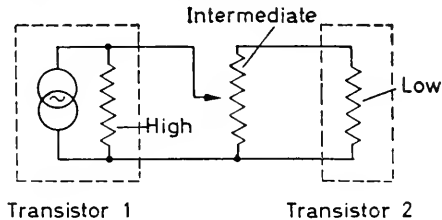


Figure 45.30

circuits the position is reversed. The output impedance of a transistor stage is larger than the input impedance of the stage which follows it, and under these conditions the arrangement of *Figure 45.29* gives poor control; a better scheme is shown in *Figure 45.30*, and this is the one usually used. It has another important advantage: the transistor is essentially a current-controlled device. It is the input current, not the input voltage, which is accurately amplified. The input resistance is that of a diode biased in the forward direction and is highly non-linear. Thus if a transistor be fed from

TRANSISTORS

a constant-voltage generator it introduces a lot of distortion, whereas if it is fed from a constant-current source it does not. With control of the *Figure 45.29* type, the source impedance seen by the load is low at low-gain settings, whereas with the *Figure 45.30* type, it is almost constant. The latter arrangement therefore gives lower distortion when the gain setting is reduced.

COMPLETE PIECES OF TRANSISTOR APPARATUS

A.c. coupled amplifiers

An audio-frequency amplifier for driving a loudspeaker, recommended by Mullard Ltd., is shown in *Figure 45.31*. Two R-C coupled stages drive a

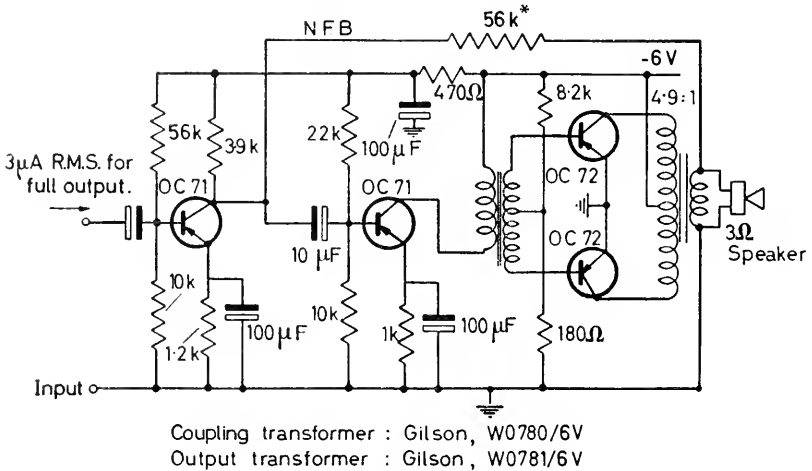


Figure 45.31

pair of output transistors in class B push-pull. The circuit is largely self-explanatory; the starred $56\text{ k}\Omega$ resistor introduces negative feedback round the output and penultimate stages. The maximum audio output power is 215 mW, and, with a 5 in. loudspeaker is sufficient to fill a large room with sound. The writer has used this amplifier, with an additional R-C coupled transistor at the front and the feedback removed, to demonstrate muscle action potentials via a concentric needle electrode.

A transistorized myographic amplifier has been described by George⁶.

For detailed design procedure for audio amplifiers, Jones and Hilbourne may be consulted⁵. Units with output powers ranging from 20 mW to 20 W are described.

Direct-coupled amplifiers

As with valves, we distinguish between 'straightforward' amplifiers and carrier amplifiers. The phototransistorized and transistorized galvanometer amplifier remains to be announced.

'Straightforward' *d.c. amplifiers*—The sensitivity of meters may be improved by the addition of a single transistor in a simple circuit such as *Figure 45.32*,

COMPLETE PIECES OF TRANSISTOR APPARATUS

due to Potok and Wales⁷. This is a voltmeter of sensitivity of $1 \text{ M}\Omega/\text{V}$. On applying the voltage to be measured, V , at the input terminals, a current V/R flows out of the base of the transistor, causing an increase in collector current $\beta(V/R)$ which is indicated by the meter. A is the main collector supply battery. In the absence of any test voltage an emitter-to-collector leakage current $(1 + \beta)I_{c(o)}$ flows, and this is 'backed-off' on the meter by current from an auxiliary battery via the associated variable

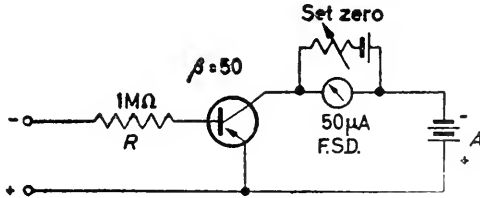


Figure 45.32

resistor, which forms the 'set-zero' control. Due to the temperature dependence of $I_{c(o)}$, a different zero setting will be frequently required. Zero-drift with temperature change can be offset by using two transistors in a push-pull circuit. Figure 45.33 shows an example due to Johnson⁸. The arrangement converts a 0–50 μA movement into an apparatus having full-scale deflection for 2 μA . Unfortunately there is nothing in this type of circuit to prevent the rise of collector current with rising temperature. Such a rise may take the transistor working points outside their working regions, so that correct amplification cannot take place.

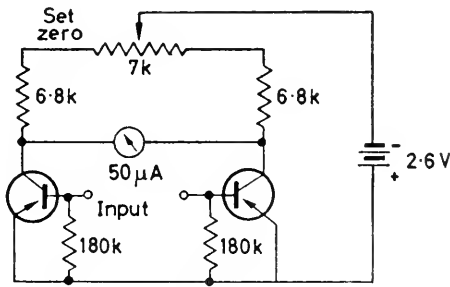


Figure 45.33

We have seen that the inclusion of resistance in the emitter circuit serves to offset the effects of temperature. We have also seen that such a resistance introduces negative feedback and reduces gain. In a.c. coupled circuits, where there is a definite minimum frequency at which the circuit is required to operate, this feedback may be eliminated by shunting the emitter resistance with a large capacitance; in direct-coupled circuits there is no such frequency and another method is necessary.

The technique is to use the transistor equivalent of the long-tailed pair in valves (Figure 45.34). This is a differential current amplifier which will discriminate against in-phase input signals. Inputs in opposite phase produce equal and opposite changes in collector current; there is no net change in the current through the common emitter resistance, and no

TRANSISTORS

feedback. The larger R_s , the better the stabilization, but to preserve a given emitter-collector voltage and current for the transistor, R_s and the positive voltage to which R_s is returned must be increased *pro rata*. Neale and Oakes have pointed out⁹ that this may lead to unnecessarily high battery voltages, and described a circuit in which an additional transistor was used to provide stabilizing action. In this manner an amplifier was constructed having a gain of 45 dB's and a zero stability corresponding to 10^{-9} amps at the input. Two transistors in push-pull earthed emitter fed a further pair in earthed collector; the latter were used to drive a meter. A fifth transistor was arranged to stabilize the first stage transistor at a low-noise working point (collector voltage between -0.2 and -1 V).

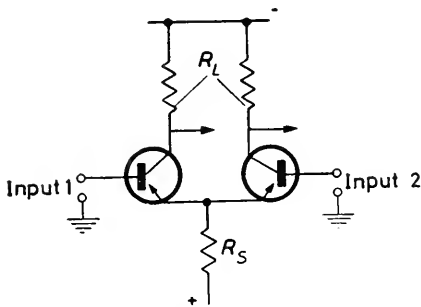


Figure 45.34

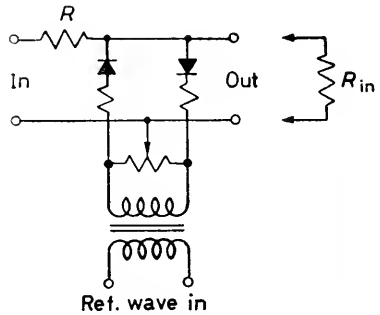


Figure 45.35

It is interesting to consider the possibility of using apparatus such as that of Neale and Oakes as a pre-amplifier for use with micro electrodes in electrophysiology. If the zero stability is 10^{-9} amps, then if the amplifier is fed from a $20\text{ M}\Omega$ microelectrode, the corresponding noise voltage at the electrode tip is 20 mV. The penetration of a cell (70 mV input or thereabouts) could probably be detected, but the device is likely to be too noisy for the extracellular recording of discharges from neighbouring cells.

Carrier amplifiers—There is no difficulty, in principle, in making a transistorized carrier amplifier merely by interposing an a.c. coupled transistor amplifier between a relay chopper and a relay phase-sensitive rectifier; yet this is seldom done. An important characteristic of transistorized devices is their small size, and there is little to be gained by using a very small amplifier if it works between two large relays. Consequently, designs for carrier amplifiers employing transistors involve chopper and rectifier devices of smaller proportions.

Some of these invoke the almost perfect performance of the recently developed silicon signal diode. We have seen in Chapter 5 how diode networks can be used for phase-sensitive rectification; the same circuits can also be used to chop. For example, the arrangement in *Figure 45.35* (which, if R is replaced by a capacitance, we have hitherto regarded as a quasi-Cowan bridge rectifier) can clearly be used as a chopper. When the reference wave transformer drives current clockwise round the diode circuit, the diodes are switched to low resistance and the output is short-circuited. When the reference wave reverses, the diode resistances become very high. The magnitude of the output depends on the input, and also on the ratio

COMPLETE PIECES OF TRANSISTOR APPARATUS

of the input resistance, R_{in} , of the amplifier the device feeds, to R . This particular circuit is therefore better for thermionic amplifiers, where R_{in} is high. It has been used with a valve amplifier by Fleming¹⁰, who reported a drift equivalent to a signal at the input of only $100 \mu\text{V}$ per hour.

A silicon diode chopper (or, more strictly, modulator) intended to feed a transistor amplifier has been described by Moody¹¹. The circuit is shown in *Figure 45.36* and is seen to resemble closely the full-wave phase-sensitive

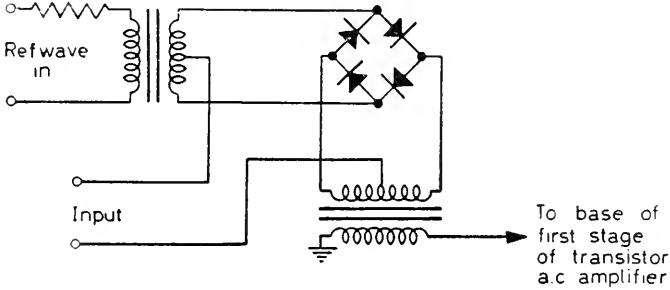


Figure 45.36

detector discussed in Chapter 6. The zero stability claimed at room temperature is equivalent to a signal at the input of less than 10^{-8} amps. It is interesting to notice that this is no better than that claimed for the 'straight-forward' d.c. amplifier of Neale and Oakes.

Transistors can be used both for phase-sensitive rectification and for chopping. A phase-sensitive rectifier has been described by Sutcliffe¹² in which two transistors act as amplifiers as well as rectifiers. A complete amplifier employing a similar transistor circuit both for chopping and rectification is due to Burton¹³. The principle of the method is illustrated

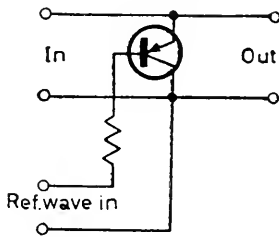


Figure 45.37

in *Figure 45.37*. We have seen that if the base of a transistor is of N germanium, both emitter and collector are of P germanium, and this suggests there ought to be some degree of reversibility between the latter two electrodes. In fact this is the case. Whether the collector be made negative with respect to the emitter (conventional usage) or whether the polarity be reversed, the current which flows between them is always large when a large current is taken out at the base, and vice versa. The only point which should be made is that, in the reversed connection, β is about one-tenth of its proper value. Thus, in *Figure 45.37*, whatever the polarity of the signal at the input, a large emitter-collector current can flow (corresponding to output

short-circuited) if current is taken out of the base by applying a reference wave which makes the base more negative than either emitter or collector. Conversely, when the reference wave goes positive, causing current to enter the base, the transistor switches to high resistance, again irrespective of the polarity of the input signal, provided the reference wave is large enough.

The transistor chopper and rectifier have respectively the simple circuits of *Figure 45.38a* and *b*. The reference wave is generated by a transistor

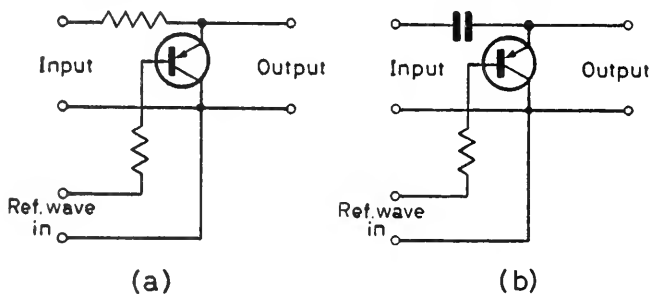


Figure 45.38

multi-vibrator (see below), the a.c. coupled amplifier employs 5 transistors, and two further transistors in earthed collector follow the rectifier and filter system to secure an output impedance of only a few tens of ohms. The apparatus is used to amplify the output of a thermocouple and, since the latter is a low resistance device, the performance is expressed by Burton in terms of voltages. The basic voltage gain of about 1,000 is reduced to 55 by negative feedback. Feedback as heavy as this enables Burton to claim a drift equivalent to only 1 mV at the input *over the ambient temperature range* -12°C to $+50^{\circ}\text{C}$. Even allowing for the fact that 1 mV across the low resistance of a thermocouple ($10\ \Omega$) represents a current change of 10^{-4} amps, the figure given augurs adequate independence of room temperature.

D.c. converters

These are devices for producing a direct output from a direct input at some other voltage. As normally used, they are arranged to produce a voltage step-up. For example, a small converter described by Johnston¹⁴ supplied an output of 30 V, 100 μA from an input at 3 V. The object was to supersede the relatively expensive HT battery in thermionic hearing aids by power derived by the converter from the relatively cheap LT cells. The efficiency of conversion was 60 per cent and the unit measured only $1\frac{3}{8} \times 1\frac{1}{16} \times \frac{1}{2}$ in. At the other end of the scale, a converter employing a power transistor¹⁵ delivered 10 kV; 100 μA for an input of 12 V, 150 mA; efficiency, 55 per cent. Much higher conversion efficiencies are possible. Perhaps the classical British paper on d.c. converters is by Light and Hooker¹⁶, but see also Light¹⁷. The simplest type is shown in *Figure 45.39*; in principle it is a transistor oscillator, with feedback achieved from collector circuit to base circuit by mutual induction, arranged so that the collector winding, L_p , forms the primary of a step-up transformer. A high alternating voltage appearing across the secondary winding L_s is half-wave rectified

COMPLETE PIECES OF TRANSISTOR APPARATUS

to produce the requisite direct output. The system is clearly reminiscent of that outlined earlier for the generation of EHT from HT by valves, but the oscillator is not of the sine wave variety. Its operation is as follows: *Figure 45.40* shows the collector characteristic for the transistor. At each cycle of oscillation the working point moves once round the locus formed by the triangular figure, *ABC*. Consider the state of affairs at point *A*; the voltage across the transistor is very low, hence nearly all the battery

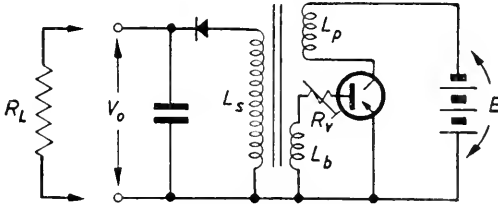


Figure 45.39

voltage appears across the primary winding L_p , and the current in it rises linearly at a rate E/L_p amps per second. If the turns ratio between L_b and L_p is $n_b : 1$ then the voltage induced across L_b by the current rising in L_p is En_b and the sense of connections is such that this drives a current $I_b = (En_b)/R_v$ out of the base of the transistor. The working point therefore moves towards *B* in *Figure 45.40*, along the characteristic corresponding to a base current of this value.

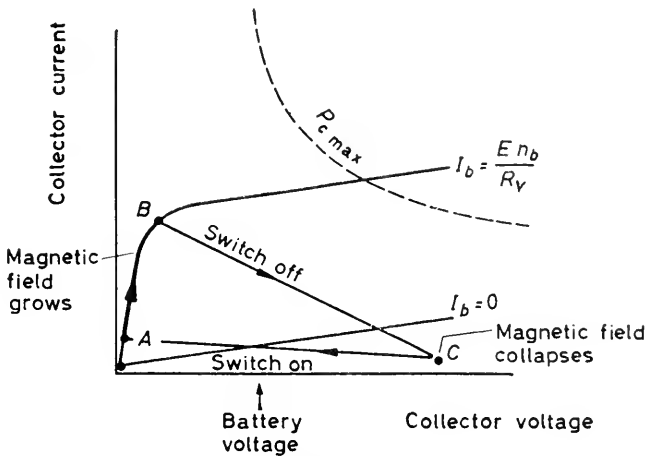


Figure 45.40

The upper limit of collector current is defined by βI_b ; that is, when the collector current reaches point *B*, it has a value of about $\beta \cdot (En_b)/R_v$ and can increase no further. The rate of rise falls away, and with it the voltage across the base winding. The base current also falls, defining a lower maximum collector current, and a cumulative action occurs as a result of which the transistor is switched to a point such as *C*. Primary current ceases abruptly. The magnetic field built up by the primary begins to collapse and in doing so does three things.

TRANSISTORS

- (1) Drives the collector *more* negative than the battery voltage.
- (2) Drives the base positive. There is then no emitter current, and only a tiny leakage current $I_{c(o)}$ from base to collector, which may be neglected.
- (3) Develops a potential difference across the secondary winding which rises rapidly until 'caught' at the steady output voltage V_o by the diode. Secondary current then flows to charge the capacitor. During this period the working point remains at C. When the energy transfer is complete,

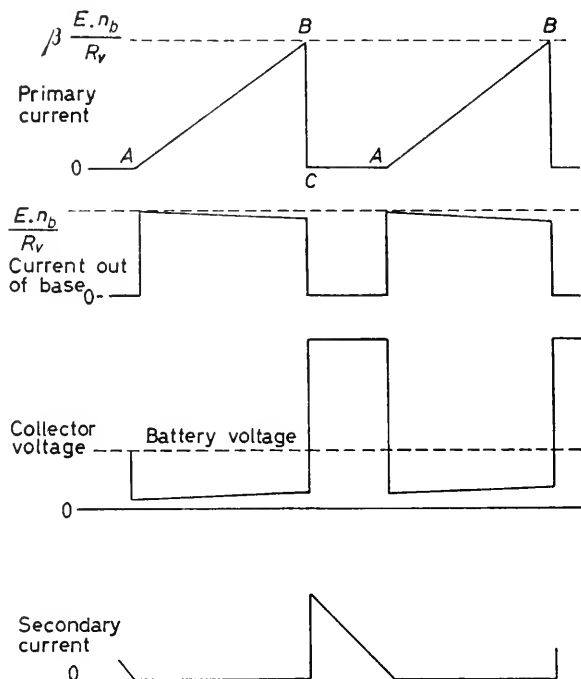


Figure 45.41

the base positivity disappears and the transistor begins to conduct from emitter to collector again. The working point reverts to *A* and the cycle repeats (Figure 45.41).

The output voltage V_o depends on the step-up ratio n_s between L_s and L_p , and on the value of the load R_L . Because the phases of magnetic field growth and collapse are quite separate in converters of this type, the growth phase is independent of the collapse phase; that is, the device draws a rather constant power from the battery, independent of the load resistance. The output voltage and current are therefore related by a rectangular hyperbola and the actual output voltage and current are determined in a particular case by the load used (Figure 45.42). The power output of the device is adjustable in practical circuits by the resistor R_f .

Notice that, except during the transition period—quickly over—between *B* and *C* in Figure 45.40, the P_{\max} hyperbola is never approached; that is, the collector dissipation is always low. This means that the efficiency is good, and that the device can handle powers much greater than the maximum

COMPLETE PIECES OF TRANSISTOR APPARATUS

permissible collector dissipation for the transistor used; but notice also that the high collector voltage to which the transistor is subjected at point *C* is related to V_o by the turns ratio n_s between L_p and L_s . Typically, a transistor might have a maximum permissible collector voltage of 30. If the battery voltage is 12, then the maximum voltage which may occur across L_p during the collapse phase is 18. Hence if $n_s = 50$, care must be taken that R_L is never so high that V_o is allowed to exceed 900 V.

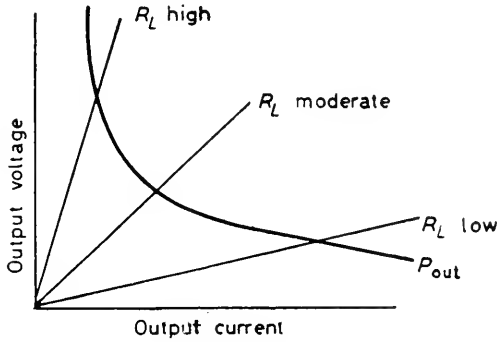


Figure 45.42

Magnetic time base

The linearly rising current in L_p (Figure 45.41) suggests that a circuit along the lines of the converter might be used as a time-base generator for a magnetically deflected cathode ray tube. In fact this is quite possible; the time-base circuit has the appearance of Figure 45.43, and has been used by the author for an experimental transistorized oscilloscope. The energy in the magnetic field following each growth has to be recovered from across

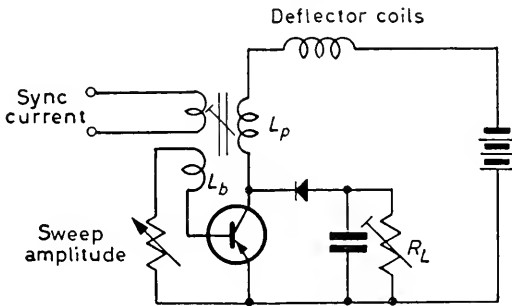


Figure 45.43

the winding L_p by the crystal diode, capacitor and load as shown. A moderately linear sweep is possible whose duration is adjusted by varying an air gap in the iron core. The difficulty is to get the flyback time T_2 in Figure 45.44 short compared with T_1 . On a repetitive time base T_2 represents 'dead' time in which events will not be drawn on the cathode ray tube face properly. It is not difficult to see that, ignoring for one moment the deflector coils, the time ratio T_1/T_2 is equal to the ratio of voltages appearing across L_p , V_2/V_1 . For the current rise to be linear, the voltage dropped across the

TRANSISTORS

transistor during the sweep (0.2 V) must be small compared with the battery voltage. A battery voltage of -6 is about the minimum. Thus we have $V_1 \approx 6$, and if the transistor has a maximum permissible collector voltage of 30, we have $V_2 = 24$. Thus $T_1/T_2 = 24/6 = 4$. This is not a very good figure for a continuous-running time base, but the position will improve as

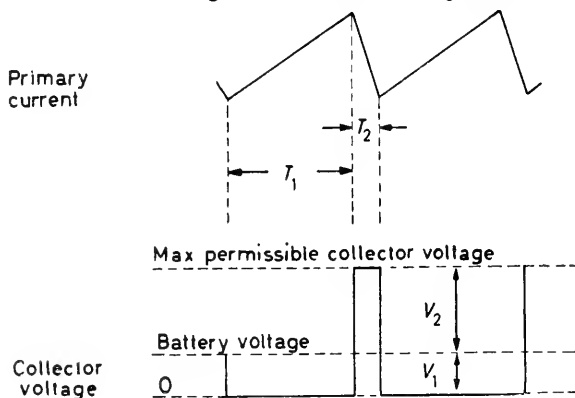


Figure 45.44

transistors having higher maximum collector voltages become available. In setting up the circuit R_L is adjusted so that the voltage across it is 30; T_1/T_2 then has its highest possible value. It is hoped in future to use the recovered energy, at present wasted in R_L , to power other transistors, or to provide HT for input cathode follower valves.

The circuit may be modified for triggered operation by the addition of a single dry cell (Figure 45.45). This drives current into the base of the transistor and the emitter-collector current has a very low value at which β is

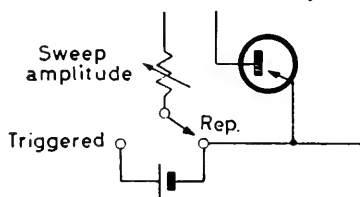


Figure 45.45

much reduced. The loop gain of the oscillator is insufficient to maintain oscillation, and the circuit gets 'stuck' at point *A* in the cycle. On the arrival of a pulse of trigger current, a voltage is induced in L_b which series-opposes the dry cell. The circuit performs one stroke, then waits for a further pulse.

It seems probable that future, more sophisticated, transistor time-bases will contain a number of stages, some transistors involved in generating accurate triangular current waves, and others acting as power amplifiers. See, for example, Nambiar and Boothroyd¹⁸.

Transistor multi-vibrators

A transistor multi-vibrator is strongly reminiscent of its valve counterpart (Figure 45.46). The subject has been discussed by Jacketts¹⁹.

Transistor R-C oscillators

A sine wave oscillator of the phase shift type is shown in *Figure 45.47*. Here again the resemblance to the corresponding valve circuit is clear. RC oscillators of this and other types have been published by Hooper and Jackets²⁰.

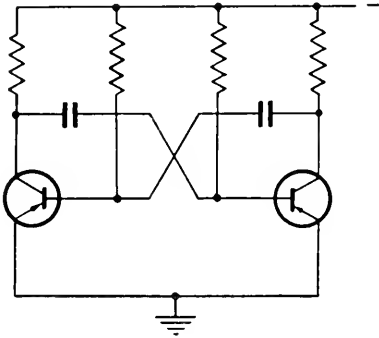


Figure 45.46

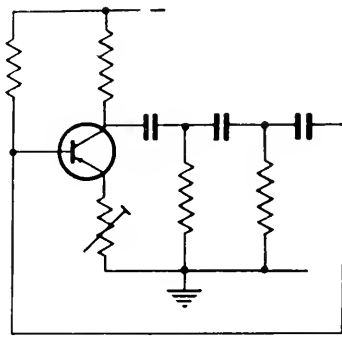


Figure 45.47

Transistor voltage stabilizers

The most elementary type of stabilizer is a single transistor connected in what is virtually the earthed collector mode (*Figure 45.48*). Such a circuit affords a measure of both forward and backward stabilization, and is the transistor analogue of the simple cathode follower arrangement in *Figure 37.11*. The output resistance in earthed collector circuits is usually given with sufficient accuracy by R_G/β , where R_G is the signal generator resistance.

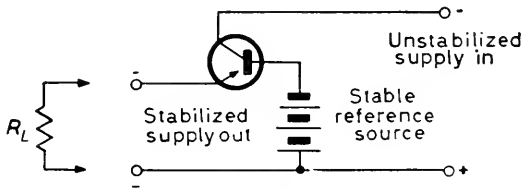


Figure 45.48

In this case R_G is the resistance of the reference battery, which should be negligible, and we have to use the exact expression which includes the usually unimportant r_b and r_e . This is

$$R_{out} = r_e + \frac{r_b + R_G}{\beta}$$

In a transistor quoted by Evans and Carroll²¹ in connection with stabilizers of this type, we have:

- $r_c = 2.5 \Omega$
- $r_b = 150 \Omega$
- $r_e = 500 \text{ k}\Omega$
- $\beta = 50$

TRANSISTORS

Taking R_G as 0, the output resistance is $2.5 \Omega + 150/50 = 5.5 \Omega$. Evans and Carroll give the forward stabilization ratio

$$F = \left[\frac{\delta V_{\text{out}}}{\delta V_{\text{in}}} \right] R_L = \text{const for the circuit as } \frac{r_b R_L}{r_c \left(R_L + r_e + \frac{r_b}{\beta} \right)}$$

If the stabilized output voltage is 6 and the current output is 60 mA, $R_L = 100 \Omega$ and $F = 1/3,300$. They point out that this simple circuit gives a satisfactory forward stabilization, but that, in that $R_{\text{out}} = 5.5 \Omega$, the backward stabilization is to be regarded as poor.

It is not necessary to use a battery as a voltage reference source. Beside the 'Accumulateur Étanche', mentioned earlier in this Part, there is an interesting device which corresponds in semi-conductor circuitry to the

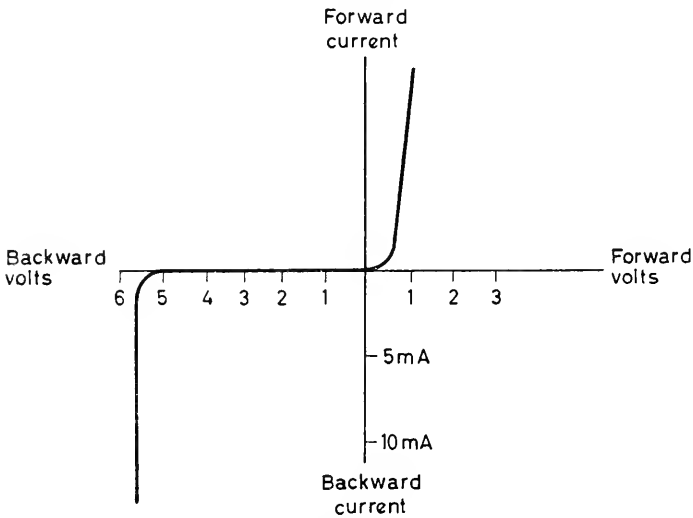


Figure 45.49

voltage reference glow discharge tube in thermionic valve practice. This is the *Zener diode*, a silicon diode having a voltage-current curve of peculiar shape (Figure 45.49). The device behaves quite normally for applied forward voltages, and for backward voltages up to a critical point. Thereafter the insulating properties suddenly break down and a reverse current can flow. When this happens the potential difference across the diode is highly independent of the reverse current. The slope of the characteristic in this region corresponds to a resistance of the order of 1 ohm only. Thus, if the reverse current through the diode can be defined at 5 mA or so by connecting it in series with an appropriate resistance to an unstabilized supply (Figure 45.50), a stable voltage, which varies according to diode type between 3 and 8, is available. The simple earthed collector stabilizer with Zener diode reference voltage has the appearance of Figure 45.51.

To improve stabilizer performance, we can emulate hard valve stabilizer practice and use an error-amplifying transistor to operate our control transistor. In Figure 45.52 a fraction of the output voltage is applied to the

COMPLETE PIECES OF TRANSISTOR APPARATUS

base of the amplifying transistor TR_1 , whose emitter is supplied with reference voltage from the Zener diode. Departures of the output voltage from the desired value alter the collector current of TR_1 and hence the

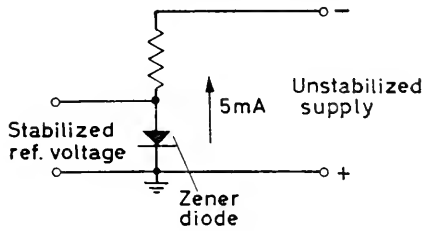


Figure 45.50

base current of TR_2 . The main output current, flowing from emitter to collector of TR_2 , is changed in such a sense as to correct the original departure.

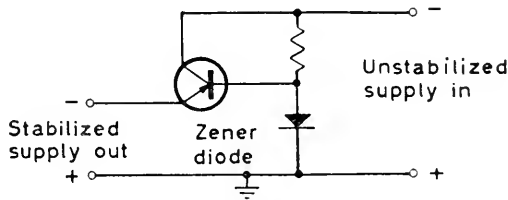


Figure 45.51

Even better performance can be secured by matching the high output impedance of TR_1 to the low input impedance at the base of TR_2 by a third transistor in earthed collector, producing an arrangement like Figure 45.53.

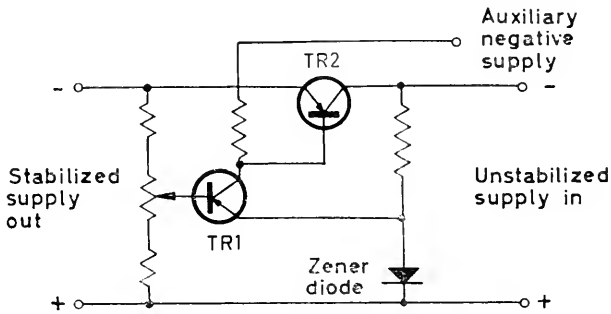


Figure 45.52

Evans and Carroll²¹ obtained a stabilized output of -6 V from an unstabilized input of -8 V and an auxiliary supply of -9 . The output resistance was less than $0.3\ \Omega$ and the stabilization ratio better than $1/20,000$.

Lloyd²² has described a transistor stabilizer which produces an output of 6 V at up to 2 amps from an input of $7.8\text{ V} \pm 10\text{ per cent}$. This would clearly be of value for supplying the heaters of valves in direct-coupled amplifiers. Five transistors are used. There is a 3 stage error amplifier, in which the transistors are in earthed emitter, whose output feeds the control transistor via a matching transistor in earthed collector. It is claimed that

TRANSISTORS

there is no detectable change in output voltage for a 10 per cent change in input voltage: nor was any ripple apparent in the output.

Brown and Stephenson²³ have published a versatile mains-driven unit delivering 0 to 30 V d.c. at up to 1 amp. This is interesting in that the error

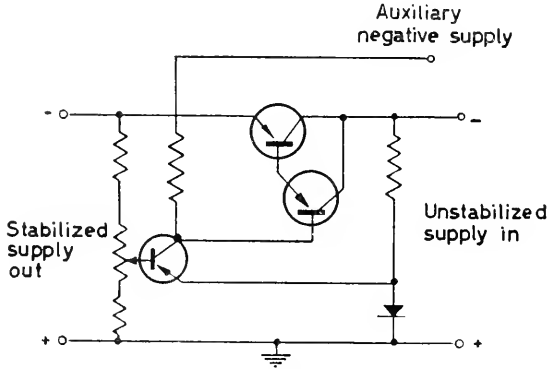


Figure 45.53

amplifier uses a long-tailed transistor pair, which should reduce the dependence of the output voltage on temperature effects. The output characteristic is given as an output resistance of less than 0.04Ω at zero frequency, and as an impedance, seen looking into the output terminals of the stabilizer, of less than 0.2Ω at all frequencies up to 100 kc/s. A 5 per cent change in mains voltage produces a 0.2 per cent change in output voltage. The ripple in the output is less than 1 mV.

Transistor current stabilizers

Power supplies delivering substantially constant currents of up to a few amperes are possible by making use of the very high collector resistance r_c of transistors in the earthed base mode. In Figure 45.54 the emitter battery

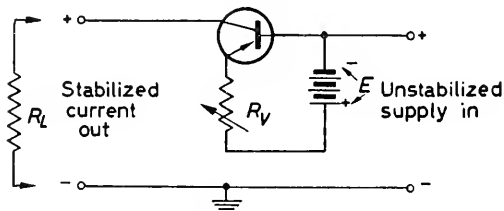


Figure 45.54

drives a current E/R_V into the base, defining a collector current $\alpha \cdot (E/R_V)$ which is highly independent of fluctuations in R_L . The relationship between possible limits of current output, of load resistance, and of input voltage used, may be derived by constructing a working region on the earthed base collector characteristic, and drawing in the load line in the normal manner.

Some other transistorized devices

A cardiometer has been described by Molyneux²⁴. To achieve a high input resistance the electrodes feed an earthed collector first stage, followed

REFERENCES

by 3 stages of amplification in earthed emitter, of which two are acceptor-tuned by negative feedback through parallel T 's. The tuned stages select the R -component of the electrocardiogram, and pass it on to a distorting transistor working with practically no forward bias, delivers a positive pulse at the collector corresponding to the peak of each R wave. This is used to trigger the transistor equivalent of a cathode-coupled flip-flop (Figure 45.55). Pulses of standard size thus derived are used to charge a

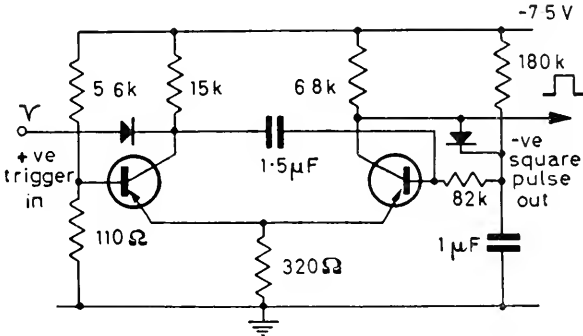


Figure 45.55

capacitor via a buffer transistor. The capacitor is discharged by a meter, whose reading is therefore proportional to the rate at which R waves arrive.

Wolfendale, Morgan and Stephenson²⁵ have reviewed the use of transistors as computing elements, and include a decade counting circuit.

Blake and Eames²⁶ have described a frequency meter for the range 0 to 100,000 counts per second. The count rate is indicated on a meter having a linear scale. The device is sensitive to input pulses only 60 mV in amplitude, corresponding to a current of 12 μ A. The accuracy claimed is ± 1 per cent.

REFERENCES

- ¹ BARDEEN, J. and BRATTAIN, W. H. *Phys. Rev.* 74 (1948) 230
- ² COCKING, W. T. *Wireless World* 61 (1955) 500
- ³ OAKES, F. *Wireless World* 61 (1955) 164
- ⁴ STUART-MONTEITH, G. *Electron. Engng.* 28 (1956) 544
- ⁵ JONES, D. D. and HILBOURNE, R. A. *Transistor A.F. Amplifiers* London; Iliffe
- ⁶ GEORGE, R. E. *Proc. E.P.T.A.* Vol 7, No. 4
- ⁷ POTOK, M. H. N. and WALES, R. A. MCF. *Electron. Engng.* 27 (1955) 344
- ⁸ JOHNSON, G. *Wireless World* 61 (1955) 31
- ⁹ NEALE, D. M. and OAKES, F. *Wireless World* 62 (1956) 530
- ¹⁰ FLEMING, L. *Electronics* 30 (1957) No. 1
- ¹¹ MOODY, N. F. *Electron. Engng.* 28 (1956) 94
- ¹² SUTCLIFFE, H. *Electron. Engng.* 29 (1957) 140
- ¹³ BURTON, P. L. *Electron. Engng.* 29 (1957) 394
- ¹⁴ JOHNSTON, D. L. *Wireless World* 60 (1954) 518
- ¹⁵ Mullard Technical Communication 2 (1956) No. 17
- ¹⁶ LIGHT, L. H. and HOOKER, P. M. *Proc. Instn. elect. Engrs.* Pt. B, 102 (1955) 775
- ¹⁷ LIGHT, L. H. *Wireless World* 61 (1955) 582
- ¹⁸ NAMBIAH, K. P. P. and BOOTHROYD, A. R. *Proc. Instn. elect. Engrs.* Pt. B, 104 (1957) 293

TRANSISTORS

- ¹⁹ JACKETS, A. E. *Electron. Engng.* 28 (1956) 184
²⁰ HOOPER, D. E. and JACKETS, A. E. *Electron. Engng.* 28 (1956) 333
²¹ EVANS, C. S. and CARROLL, J. L. *Electron. Engng.* 29 (1957) 143
²² LLOYD, M. A. *Electron. Engng.* 29 (1957) 197
²³ BROWN, T. H. and STEPHENSON, W. L. *Electron. Engng.* 29 (1957) 425
²⁴ MOLYNEUX, L. *Electron. Engng.* 29 (1957) 125
²⁵ WOLFENDALE, E., MORGAN, L. P. and STEPHENSON, W. L. *Electron. Engng.*
29 (1957) 136
²⁶ BLAKE, L. R. and EAMES, A. R. *Electron. Engng.* 28 (1956) 322

INDEX

- A.c. relays 522
Acceptor amplifier 198, 200, 203, 631, 709
Accumulator cells 288, 660
Active elements 3
Air gap 54
Amplification factor 135
Amplifier, acceptor 198, 200, 203
— carrier amplifier 619
— differential power 193
— differential voltage 181
— double sided 176
— galvanometer amplifier 620
— loudspeaker 636
— pentode, simple power 147
— — simple voltage 147
— practical a.c. coupled 618, 619, 634
— practical chopper amplifier 632
— practical direct coupled 623, 634
— transistor a.c. coupled 684, 696
— transistor chopper 698
— transistor direct coupled 697
— triode, simple power 140, 141
— — simple voltage 137
— tuned 196
Anode 118, 133
— characteristic of tetrode 145
— characteristic of triode 135
— incremental resistance 135
Astigmatism 455
Attenuator
— bridged T 22
— constant resistance 16
Autotransformer 306
Avalanche formation in G-M tubes 426

Balance
— control in differential voltage amplifier, gain 189
— control in differential voltage amplifier, static 190
Balanced input 181
Ballast resistance 119
Base (of transistor) 678
Batteries 283
Bi-stable circuit 124, 238

Bleater 636
Blocking capacitance 153,
— of amplifiers 616
— oscillator 215
Blumlein integrator 241
Bridge
— Cowan 117
— for use with mechanoelectric transducer 491
— rectifying 97, 315, 317
— unbalanced Wheatstone 12
— Wien 48
Brimistor 340
Bucket machine 498

Cables 325
Cadmium selenide photocell 377
Calibrators 628
Capacitance
— in parallel 25
— in series 26
— manometer 505
— single 25
Capacitors 297
— electrolytic 302
— high k ceramic 301
— low k ceramic 300
— mica 299
— paper 298
— variable 303
Carbon resistor noise 254, 293
Cascade circuit 257, 590
Catching 109
Cathodally screened cathode follower 259
Cathode
— cold 118
— coupled flip-flop 235
— directly heated 132
— follower 169, 259
— indirectly heated 133
— of photoemissive cell 356
— ray oscilloscope light sources 352
— — tube, electrostatic type 452
— — — hybrid type 469
— — — magnetic type 460
Chassis 327

INDEX

- Child's law 134
- Chokes
 - design of small air cored 308
 - — — — iron cored 309
 - smoothing 82, 104, 308
- Chopper amplifier 619, 628, 632
- Clamping 112
- Class B operation 180
 - C operation 210
- Climbing amplifier 149
- Clipping 110
- Collector 678
- Colour code
 - — ceramic capacitors 301
 - — mica capacitors 300
 - — resistors 291
- Colour temperature 338
- Concertina phase splitter 168
- Connecting wire 325
- Connectors 322
- Constant voltage transformer 307
- Control ratio 128
- Copper oxide rectifiers 316, 317
 - — signal diodes 318
- Cossor double beam C.R.T. 183, 456
- Coulomb definition 6
 - in capacitors 25
- Counters
 - cathode ray 649
 - drop 646
 - glow-discharge 649
 - Post Office, electromagnetic 531
 - radioactive, errors of 431
 - relay scale-of-two 527
- Coupling
 - a.c., methods 153
 - — transistors 686, 687
 - battery 150
 - direct, methods 149
 - factor, k 61
 - interstage 149
 - neon lamp 152
 - of mutual inductance 60
- Cowan bridge 117
- Critical damping 72, 73, 76, 77, 81
 - inductance 105
- Crossover distortion 181
- Crystal
 - piezo-electric, transducers 481, 492
 - quartz, oscillator 217
 - quartz, timer 644
- D.c. converters 700
 - restoring 112
- Decibel 21
- Decoupling 160, 688
- Deflection
 - amplifiers for electrostatic C.R.T. 458
 - amplifiers for magnetic C.R.T. 466
 - defocusing 463
 - sensitivity 454
- Dekatron counter 649, 650
- Desynn synchronous link 496
- Difference diode 124
- Differential
 - cathode resistor 191
 - power amplifier 193
 - voltage amplifier 181
- Differentiating circuit 237
- Diode 96
 - hard 134
 - intervalve coupling by soft cold cathode 152
 - non-thermionic 315
 - signal 317
 - soft, cold cathode 118
 - — difference 124
 - — hot cathode 126
 - Zener 706
- Discrimination ratio 183
- Distortion
 - crossover 181
 - harmonic, in power stages 178
 - reduction of, by negative feedback 164
- Drift 161
- Dry Leclanché cells 283
- Duddell oscillograph 452
- Dushman's equation 118
- Dynamic resistance 78
- Eccles-Jordan circuit 238, 647
- Efficiency, power transformer 70
 - triode power amplifier 142
- Electrodes
 - concentric 576
 - fixation of 572
 - implanted 570
 - metal micro 574
 - microcapillary 534
 - non-polarizable 568
 - oxygen tension 576
 - pH 578
 - simple metal 570
- Electrodynamic transducers 482, 492
- Electroluminescent light sources 354
- Electron multiplier photocells 361
- Electronic test equipment for design and maintenance work 673

INDEX

- Emitter (of transistor) 678
 Equivalent circuit
 — — barrier-layer cell 368
 — — demonstrating pick-up of interference 656 etc.
 — — demonstrating stimulus artefact 609 etc.
 — — earthed-base transistor 692
 — — earthed-emitter transistor 682, 691
 — — force-recording apparatus 475, 477
 — — input of flip-flop 443
 — — intervalve couplings 153, 155
 — — inter-transistor coupling 687
 — — microelectrode 552, 554, 559,
 — — multi-vibrator 230
 — — pen recorder 495
 — — pentode amplifier 148
 — — piezo-electric transducer 481
 — — transmission line in mechanical systems 478
 — — triode amplifier 140
 Exponential response 28, 59, 552
 Extra high tension (EHT) 284, 363, 424, 457, 597

 Farad 25
 Faults
 — in components 665, 666
 — in equipment, complete 665
 — in equipment, illusory 669
 — in equipment, intermittent 668
 — in equipment, partial 667
 Feedback, negative current 250
 — negative voltage 164
 Ferroxcube 309
 Filter
 — classical 83
 — other LC types 91
 — R-C band-pass 46
 — — cascaded sections 39
 — — fed from real generator 38
 — — high-pass 35
 — — low-pass 35
 — — with limited phase shift 43
 Flicker noise 255
 Flip-flop 235
 — transistorized 709
 Flux, magnetic 54
 Frequency response
 See under Steady-state response or transmission characteristic
 Fuses 324

 Gain
 — balance control for differential voltage amplifiers 189
 — control of transistor circuits 695
 — control for differential voltage amplifiers 185
 — loop 165
 — of earthed emitter transistor 683
 — stabilization of, by negative feedback 167
 — voltage amplifier, pentode 147
 — — — triode 138
 Galvanometer amplifier 620
 Gas discharge lamps 344
 — chamber counting 423
 — filled photocells 358
 Geiger tubes 424, 426, 435
 Generators
 — constant current 3, 134, 148
 — constant voltage 3
 — ideal 3
 — real 8
 Germanium
 — junction photodiode 378
 — junction rectifiers 316
 — junction transistor 677
 — signal diodes 317
 Glow modulator tube 351
 Grid 128, 133
 — bias 135
 — — automatic 154
 — control 144
 — current control 624
 — screen 144
 — suppressor 146

 Hanna design procedure 313
 Hard valves 132
 Heat 383
 Heater power supplies for valves 284, 288, 594
 Henry 56
 High-speed relay 519
 High tension (HT) 134, 137
 — — batteries 284
 — — supplies, mains derived, stabilized 586, 587
 — — supplies, mains derived, unstabilized 585
 Homogeneous resistance noise 253
 Humidity control 416
 — measurement 413

 Ideal generators 3
 — transformer 64

INDEX

- Impedance 32
 - mechanical, of transducers 478
 - of parallel resonant circuit 79
 - of series resonant circuit 76
 - terminal, of classical low-pass T section 86
 - terminal, of classical high-pass T section 88
 - terminal, of classical low-pass π section 88
 - terminal, of classical high-pass π section 90
- Indicator lamps 324
- Inductance
 - leakage 65
 - mutual 56
 - optimum 104
 - primary 65
 - self 56
 - — in parallel 57
 - — in series 57
- Induction coil stimulator 602
- Input
 - capacitance of triode valve 144
 - impedance of cathode follower 169
 - resistance 135
 - — of earthed-collector transistor 695
 - — of earthed-emitter transistor 683
- Integrator, Blumlein or Miller 241
- Interference
 - compensation for 664
 - fields, electric 656, 659, 660, 661
 - — magnetic, 657, 659, 661, 663
 - — radiated 658, 660
 - identification of 663
- Interstage coupling 149, 686, 687
- Ion burn 465
 - trap 465
- Ionization chambers 425
 - potential 118
- Ionizing radiation, detection of 422
- Isotope characteristics 419

- j operator 33
- Johnson noise 253
- Joule 6
- Junction transistor 677

- Kalium cells 285
- Knobs 322

- Lead-acid cell 288
- Lead-sulphide photoconductive cell 373
- Light sources
 - — cathode ray oscilloscope 352
 - — electroluminescent 354
 - — gas discharge lamps 344
 - — linear 351
 - — mercury vapour lamps 344
 - — tungsten lamps 339
 - — xenon arc lamps 348
 - — zirconium arc lamps 350
- Linear light sources 351
- Load 3
 - line 139, 142, 685
 - matched 9
 - polarization 162
- Loaded RC filter 37
- Long-tailed pair 181, 697
- Loop gain 165
- Loudspeaker amplifier 636

- M motor 496
- Magnetic circuit 54
 - tape recording methods 641
- Magneto-motive force 54
- Magslip 497
- Mallory cell 285
- Manometer, capacitance 505
- Mark-space ratio 226
 - — control of temperature 398
- Matthews oscillograph 452
- Maximum power transfer theorem 9
- Mechanical impedance of transducers 478
 - properties of transducers 472
- Mercury vapour lamps 344
- Meters
 - electrostatic 446
 - moving-coil 448
 - moving-iron 447
- Microcapillary electrodes
 - — connection to 544
 - — construction 535, 560
 - — dimensions 542
 - — groups 561
 - — measurements with 547
 - — methods of filling 541, 558
 - — polarization of 559
 - — resistance measurement 543
 - — screening of 545
- Microphony 256
- Miller effect 144
 - run-down 241
 - transition 246

INDEX

- Modulus 31
- Motor boating 159
- Multi-vibrator
 - cathode coupled 231
 - symmetrical 227
 - transistorized 704
 - used as divider 645
 - used as stimulator 604
- Mutual conductance 135
 - inductance 56
 - resistance 692

- Negative current feedback 250
 - feedback and noise 256
 - resistance 170
 - voltage feedback 164
- Neon lamp coupling 152
 - stimulator 603
- Networks, easily solved 9
 - T 17
 - π 17
- Ni-Fe cells 289
- Noise 253
 - as affecting amplifier gain 616
 - carbon resistor 254
 - flicker 255
 - homogeneous resistance 253
 - in variable resistors 296
 - negative feedback and 256
 - partition 255
 - shot 254
- Nomotron counter 649
- Non-polarizable electrodes 559, 568
- Non-thermionic diodes 315
- Null indicators, magic eye 451
 - neon lamp 451

- Ohm's law 6
 - — 'magnetic' 54
 - — with reactances present 32
- Optimum inductance 104
 - load 142, 147
- Oscillator
 - blocking 215
 - crystal 217
 - LC 208
 - RC 221
 - — transistorized 705
 - relaxation 125
 - sine wave 208
 - square wave 226

- Output
 - impedance of cathode follower 173
 - power of triode valve 142
 - resistance of earthed-collector transistor 695
 - resistance of earthed-emitter transistor 683

- Panel 327
- Parallel
 - capacitances in 25
 - resistance and capacitance 40, 41
 - resistances in 7
 - resonance 78
 - self inductances in 57
 - T filter 50
- Partition noise 255
- Passive elements 3
- Peak inverse voltage (P.I.V.) 315
- Penthouse waveform 251
- Pentode valve 146
- Penwriter 494, 638
 - amplifier 637
- Permeability 54
- Phantastron 245
- Phase
 - angle for series C and R 31
 - of capacitance current 27
 - sensitive detectors 114
 - shift, classical filter sections, low-pass 86
 - — — — high-pass 88
 - — relationship to transmission characteristic 52
 - — RC filters, limited 43
 - — using j operator 36
 - shifting circuits 45
 - splitter, concertina 168
- Photoelectric detectors 356
 - — cadmium-selenide photocell 377
 - — electron-multiplier photocells 361
 - — germanium junction photodiode 378
 - — lead-sulphide photoconductive cell 373
 - — phototransistor 381
 - — selenium barrier layer photocell 367
 - — types of cathode 356, 357, 358
 - — vacuum and gas filled photocells 358
- Photographic recording methods 639
- Photometric units 333

INDEX

- Phototransistor 381
- Piezo-electric transducers 481, 492
- Pin-cushion distortion 455, 463
- Point contact transistor 677
- Polarization of electrodes 559, 568
 - of load 162
- Polarized relay 520
- Post deflection acceleration (P.D.A.) 459
- Potential divider
 - — coupling method 151, 156
 - — made of resistances 7
 - — partially reactive 34
- Potentiometer 7
 - cam-corrected 485
 - control of power 13
 - feeding screened cable 42
 - liquid 486
- Power
 - amplifier, differential 193
 - amplifier, simple 140, 141
 - in capacitance 27
 - in inductance 58
 - supplies, cold cathode stabilized HT 586
 - — electronically stabilized HT 587
 - — for EHT 597
 - — for valve heaters 594
 - — stabilized current 600
 - — — mains 600
 - transfer theorem, maximum 9
 - transformer 65, 68
- Preferred value system 292
- Proportional counters 438
- Pulse, square, generating circuits 235
 - transformer 65, 66
 - triangular, generating circuits 239
- Push-pull
 - power stage 177, 696
 - — for loads without a centre tap 194
 - voltage amplifier 176
- Q* 75, 79
 - of acceptor amplifiers 199, 202, 205
- Quartz crystal oscillator 217
 - crystal timer 644
- Rack 327
- Radioactive counting errors 431
- Rationalization 36
- Reactance capacitive 27
 - inductive 58
- Reactivation of dry cells 285
- Real generators 8
 - transformer 64
- Rectification
 - full-wave 97
 - half-wave 97
 - shunt diode 114
 - smoothed 99
- Rectifier 316
- Reference wave 115
 - tube, voltage 119
- Reflection, through real transformer 64
- Regulation of power supply 101
- Rejection ratio 183
- Relaxation oscillator, difference diode 125
 - — thyatron 129
- Relay
 - circuits 524
 - high speed 519
 - notation 510
 - others 521
 - P.O. 3000 type 512
 - polarized 520
- Reluctance 54
- Resistance capacitance
 - — band-pass filter 46
 - — coupling between valves 153
 - — filters with cascaded sections 36
 - — — limited phase shift 43
 - — high-pass filter 35
 - — low-pass filter 35
- Resistance thermometer 389
- Resistors 291
 - carbon composition 291
 - high stability 293
 - variable 294
 - wirewound 294
- Resonance
 - in signal transformers 81
 - parallel 78
 - series 74
- Rheostat 13
- Richardson's equation 118, 127
- Ripple estimation 102
- R.M.S. values 23
- Running voltage 119
- Sanatron 244
- Saturated diode 134
- Saturation
 - due to load polarization 162
 - in swinging choke 104
 - of transformer core 70
- Scalers 527, 647

INDEX

- Scintillation counters 439
 — counting 423
 Screen danger! red hot 215
 — grid 144
 Selenium barrier layer photocells 367
 Selenium-iron rectifier 316, 317
 — signal diodes 318
 Selsyn 497
 Series
 — capacitances in 26
 — resistances in 7
 — — and capacitance in 27
 — — and self inductance in 58, 59
 — — capacitance and inductance in 72
 — resonance 74
 — self inductance in 57
 Shot noise 254
 Shunt 7, 448
 Signal 4
 — transformer 65, 68
 Silicon junction rectifiers 316
 — junction signal diodes 318
 Sine-wave oscillators 208
 Single-sided amplifiers 159
 Silver-zinc cells 290
 Slow-motion drive 324
 Smoothing section 82, 585
 Soft valves 118
 Space charge 126, 148
 — — limited 134
 Split-field motor 499
 Square wave oscillators 226
 Squegging 216
 Stabilization
 — backward 122
 — forward 122
 — of gain by negative feedback 167
 Stabilized EHT supplies 599
 — heater supplies 595
 — HT supplies 586, 587
 — supplies using transistors 705, 708
 Stabilizer tube 121, 320
 Star-delta transformation 11
 Steady-state response 5
 — — acceptor amplifiers 198, 201, 205
 — — cascaded RC sections 40
 — — cathodally screened cathode follower 259
 — — intervalve couplings 154
 — — limited phase-shift filters 44
 — — parallel T filter 51
 — — penwriter 495
 — — RC band-pass filter 48
 — — RC filters 35, 40
 — — rejector amplifier 207
 — — series L.C.R. 75
 — — signal transformer 69
 Step function 4
 Stimulators
 — induction coil 602
 — more elaborate types 605, 606, 608
 — multi-vibrator 604
 — neon lamp 603
 — R.F. coupled 613
 Stimulus artefact, discussion on avoidance of 607
 Stopper resistors 601
 Striking voltage 118
 Swinging choke 104
 Switches 321
 Synchro 497
 Tagboards and tagstrips 322
 Tape recording methods 641
 Tapered sections 40
 Teledeltos paper 638
 Temperature 383
 — control 393
 — measurement 385
 Test equipment for design and maintenance work 673
 Tetrode
 — kink 145
 — valve 143
 — valve, beam 147
 Thermal noise 253
 — relay 523
 Thermionic emission 118, 126, 132
 Thermistor 223, 390, 392, 411
 Thermocouple 386
 Thévenin's theorem 12
 Thyatron 128
 — switch 130
 Time constant of C and R 29
 — constant of intervalve couplings 158
 — constant of L and R 59
 Timer unit, generating 10 msec pips 644
 Tools for electronic work 329
 Transducers
 — bucket machine 498
 — circuits for use with 500
 — — Velodyne 504
 — classification of 480
 — Desynn 496
 — electrodynamic 482, 492
 — M motor 496
 — magflip, synchro or selsyn 497
 — mechanical impedance of 478

INDEX

- Transducers—*contd.*
 — mechanical properties of 472
 — miscellaneous 490
 — piezo-electric 481, 492
 — split-field motor 499
 — synchronous links 496
 — variable capacitance 489
 — — inductance 488
 — — resistance 485
 Transformer
 — action, basis of 63
 — autotransformer 306
 — constant voltage 307
 — design of small iron-cored 309
 — ideal 64
 — mains 305
 — power 65, 68
 — pulse 65, 66
 — real 64
 — reflection through 64
 — signal 65, 68
 — variac 307
 Transient response 5
 — — cathodally screened cathode follower 261
 — — intervalve couplings 158
 — — parallel L, C and R 77
 — — parallel R and C 41
 — — penwriter 495
 — — pulse transformer 67
 — — series L, C and R 73
 — — series R and C 27
 — — series R and L 58
 Transistor
 — current stabilizer 708
 — flip-flop 709
 — junction 677
 — multi-vibrator 704
 — point contact 677
 — R-C oscillator 705
 — time base 703
 Transitron 236
 — Miller 246
 Transmission characteristic 35
 — — classical high-pass section 88
 — — — low-pass section 85
 — — smoothing section 82
 — — relationship between, and phase shift 52
 — factor of attenuator 17
 — — of potential divider 7
 Trapezium distortion 455
 Triangular current wave generators 247
 — voltage wave generators 239
 Trigger tube 127, 652
 Triggered pulse generators 235
 Triode 135
 Tubes *see* Valves
 Tuned amplifiers 196
 Tungsten lamps 339
 Uniselectors 531
 Vacuum photocells 358
 Valves (Tubes)
 — beam tetrode 147
 — difference diode 124
 — hard 132
 — hard diode 134
 — hot cathode diode 126
 — parameters 135
 — pentode 146
 — practical choice of 319
 — primed stabilizer 128
 — reference 119
 — soft 118
 — stabilizer 121
 — tetrode 143
 — thyatron 128
 — trigger 127
 — triode 135
 — variable μ pentode 222
 Valveholders 321
 Variac 307
 Velodyne 193, 504, 600
 Vertical fall 241
 Vibrating capacitor amplifier 619, 633
 Voltage
 — amplifier, differential 181
 — — pentode simple 147
 — — triode simple 137
 — multiplier rectifying circuits 106
 — rating of resistors 293, 294, 296
 — reference tubes 119
 — running 119
 — stabilizer tubes 121
 — striking 118
 Watt 6
 Wien bridge 48
 Working region
 — — pentode 147
 — — transistor 684
 — — triode 138
 Workshop facilities 329
 Xenon arc lamps 348
 Zener diode 706
 Zirconium arc lamps 350







