

FILE 60W

LAMONT-DOHERTY GEOLOGICAL OBSERVATORY
OF COLUMBIA UNIVERSITY

TREE-RING LABORATORY
Palisades, New York 10964

*

Technical Report No. CU-1-81/TR1

This material is based upon work supported by the
National Science Foundation under Grant No. ATM79-19223,
Gordon C. Jacoby, Principal Investigator.

*

THE CUBIC SMOOTHING SPLINE AS A DIGITAL FILTER

by

Kenneth Peters and Edward R. Cook

October 1981

*

Reproduction in whole or in part is permitted for any purpose of the
United States Government. Others must secure the author's
permission for use of this material.

*

Distribution of this document is unlimited.

LAMONT-DOHERTY GEOLOGICAL OBSERVATORY
of Columbia University
TREE-RING LABORATORY
Palisades, New York 10964

*

Technical Report No. CU-1-81/TR1

*

This material is based upon work supported by the
National Science Foundation under Grant No. ATM79-19223,
Gordon C. Jacoby, Principal Investigator.

*

THE CUBIC SMOOTHING SPLINE AS A DIGITAL FILTER

by

Kenneth Peters and Edward R. Cook
October 1981

*

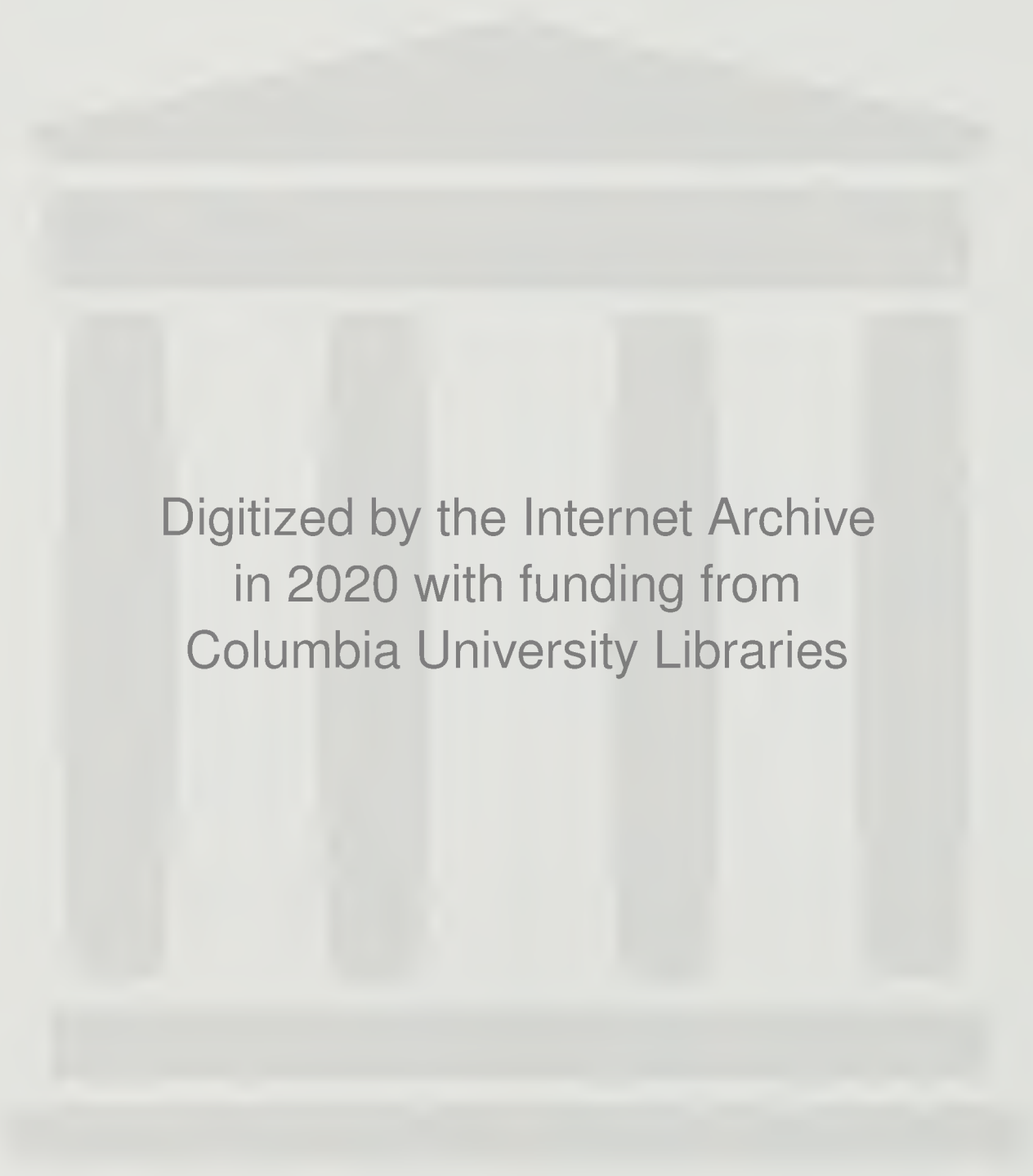
Reproduction in whole or in part is permitted for any purpose of
the United States Government. Others must secure the author's
permission for use of this material.

*

Distribution of this document is unlimited.

*

Any opinions, findings, and conclusions or recommendations expressed
in this publication are those of the authors and do not necessarily
reflect the views of the National Science Foundation.



Digitized by the Internet Archive
in 2020 with funding from
Columbia University Libraries

<https://archive.org/details/cubicsmoothingsp00pete>

THE CUBIC SMOOTHING SPLINE AS A DIGITAL FILTER

Kenneth Peters and Edward R. Cook

ABSTRACT

Reinsch (1967) introduced the cubic smoothing spline as the solution to a certain variational problem in the theory of data smoothing. This smoothing operator has good transient response, a minimal curvature property, and compensation for end effects so that all the data is used. In addition, the data can have arbitrary spacing and weighting.

In this paper we represent Reinsch's equations in terms of convolutions for the special case of equally-spaced and equally-weighted data. We then use a Fourier transform to obtain a frequency domain representation, solve the characteristic polynomial of the system, and apply an inverse transform to the frequency response function to obtain the impulse response function. The Lagrangian multiplier of the variational problem parameterizes the family of digital filters derived in this manner. This representation which is exact for finite data sets extends the range of application of the smoothing spline and, incidentally, simplifies the algebra and reduces the amount of computer time and storage required, by allowing one to select the degree of smoothing on the basis of frequency domain considerations. One can do time domain filtering without end effects in a computationally efficient manner.

Smoothings of four tree-ring width sequences from forest interior sites are presented to illustrate the particular suitability of these filters for data exhibiting both non-stationary and episodic behavior.

THE CUBIC SMOOTHING SPLINE AS A DIGITAL FILTER

I INTRODUCTION

Reinsch [7] discusses the problem of smoothing noisy data when the underlying function is not known. We begin with a brief statement of his main results in his own notation and terminology. He suggests that the fitting function, $g(x)$, minimize the total square curvature,

$$\int_{x_0}^{x_n} \left[\frac{d^2g(x)}{dx^2} \right]^2 dx \quad (1a)$$

under the constraint that a weighted sum of squares of the residuals be bounded above,

$$\sum_{i=0}^n \left[\frac{g(x_i) - y_i}{\delta y_i} \right]^2 \leq S; \quad (1b)$$

where the y_i represent the noisy data, the δy_i are a set of weights and S is a scaling parameter. Equivalently, introducing a Lagrangian multiplier, p , and an auxiliary variable, z , the functional

$$\int_{x_0}^{x_n} \left[\frac{d^2g(x)}{dx^2} \right]^2 dx + p \left\{ \sum_{i=0}^n \left[\frac{g(x_i) - y_i}{\delta y_i} \right]^2 + z^2 - S \right\} \quad (1c)$$

is to be minimized.

The solution is a cubic spline, i.e., a concatenation of cubic segments:

$$g_i(x) = a_i + b_i (x-x_i) + c_i (x-x_i)^2 + d_i (x-x_i)^3, \quad x_i \leq x < x_{i+1}, \quad (2)$$

such that $g(x)$, $dg(x)/dx$, and $d^2g(x)/dx^2$ are continuous at x_i for $i=1, n-1$.

If the Lagrangian multiplier, p , is known, the values c_i can be computed from

$$(\hat{Q}D^2Q + pT)c = p\hat{Q}y, \quad (3)$$

where the circumflex denotes transposition, and the a_i from

$$a = y - p^{-1}D^2Qc. \quad (4)$$

Then

$$\hat{d}_i = (c_{i+1} - c_i) / 3h_i, \quad i = 0, n - 1 \quad (5)$$

where, for the sake of uniformity of notation, Reinsch chooses to set c_0 and c_n equal to 0; and

$$b_i = [(a_{i+1} - a_i)/h_i] - c_i h_i - d_i h_i^2, \quad i = 0, n - 1. \quad (6)$$

The notation used is Reinsch's:

$$h_i = x_{i+1} - x_i; \quad (7a)$$

$$\hat{c} = (c_1, \dots, c_{n-1}); \quad (7b)$$

$$\hat{y} = (y_0, y_1, \dots, y_{n-1}, y_n); \quad (7c)$$

$$\hat{a} = (a_0, a_1, \dots, a_{n-1}, a_n); \quad (7d)$$

$$D = \text{diag} (\delta y_0, \dots, \delta y_n); \quad (7e)$$

$$T = \text{a positive-definite tridiagonal matrix of order } n-1, \quad (7f)$$

$$\text{where } t_{ii} = 2(h_{i-1} + h_i)/3, \text{ and } t_{i,i+1} = h_i/3 = t_{i+1,i};$$

$$Q = \text{a tridiagonal matrix with } n+1 \text{ rows and } n-1 \text{ columns} \quad (7g)$$

$$\text{where } q_{i-1,i} = 1/h_{i-1}, \quad q_{ii} = - (1/h_{i-1} + 1/h_i), \text{ and}$$

$$q_{i+1,i} = 1/h_i.$$

For Reinsch's intended application to data smoothing it is natural to specify the weights δy_i , usually as the square root of the error variance at $x=x_i$ if it is known or can be estimated, and S , as the value of a chi-square variable with n degrees of freedom. He then proceeds to

and may be of interest to the reader. Reinsch [8] recasts the problem of reference [7] in much more general terms and shows that similar results hold. Demmler and Reinsch [4] achieve by advanced methods a very general result of which formula (16) of this paper is a special case. Wahba [9] discusses the problem of optimizing the degree of smoothing given certain assumptions regarding the random component of the data and gives explicit formulas which can be used in conjunction with the heuristic considerations which motivated this paper. Wahba and Wold [10] and Craven and Wahba [3] apply a method of cross-validation to select the smoothing parameter and in [11] Wahba and Wold discuss the use of periodic splines in spectral density estimation which is related to a suggestion for natural splines in this paper in Section V. Wold [12] is concerned with the suitability of splines of various kinds from a data analytic point of view. In this paper we are concerned with a discrete problem -- the relationship between the vectors y and a . In the limit of large p there are no digital filtering effects. Horowitz [6] considers the continuous problem of the filtering effects of common interpolating splines. His analysis may be of interest if the discrete data arise from accumulation or from sampling continuous records. The references given here should provide an adequate introduction to smoothing splines.

The rest of the paper consists of four sections and two appendices. Section II contains the derivation of the convolutional representation and of the frequency response function. Section III contains the solution of the characteristic equation of the system. Section IV contains the derivation of the impulse response function. Section V includes a discussion of how to interpret and apply the representation derived in sections II, III and IV, and an example which illustrates the time domain behavior of

the cubic smoothing spline. Appendix A is a discussion of this behavior in the limit of small curvature, and Appendix B contains an analysis of the coupled ordinary differential equations of which the cubic smoothing spline is an interpolated finite difference solution.

The reader interested only in practical applications should read sections I and V and consider the main practical results which are in equations (16) and (33) and figures 1 and 5. The critical reader will want to include the beginning of section II. The appendices and the parts of the text relating to the correction terms (indicated by primes) and the small curvature limit, can be ignored, but clues to understanding the Fourier representation are contained therein. Finally, those filters generated by the real negative roots of the characteristic polynomial are of little practical interest since they result in very little smoothing.

II THE CONVOLUTION REPRESENTATION AND THE FREQUENCY RESPONSE FUNCTIONS

Consider the special case of equations (3) and (4) for which $h_i = 1.0 = \delta y_i$ for all i . In this case the matrices Q and T and $R = \hat{Q}Q + pT$ will be Toeplitz, i.e., all the terms on the same diagonal will be equal, and symmetric. The matrix-vector product Rc will be equivalent to convolution with a symmetric (phase-free) filter except for the first and the last two points. In fact these restrictions are necessary because convolution is defined as a time-invariant linear process, and it is impractical to do Fourier transforms on unequally spaced data. The particular numerical values of the spacing, in this case unity, serves to scale or normalize the parameter p which has the units of x^{-3} . With these restrictions then the desired representation can be achieved as follows.

First, extend a typical row of the matrix R by appending an infinite number of zeroes on each end, and do the same for Q and \hat{Q} . This extension is necessary because the limits on the Fourier integral are infinite. The particular form of the extension (zeroes) is dictated by the band form of the corresponding matrices. Thus define

$$r_p = \{\dots, 0, 1, (-4+p/3), (6+4p/3), (-4+p/3), 1, 0, \dots\} \quad (8)$$

and

$$q = \{\dots, 0, 1, -2, 1, 0, \dots\} \quad (9)$$

in accordance with (7f) and (7g). The subscript p makes the parameterization explicit. The vectors c, y and a can be extended in a similar manner, i.e.,

$$c = \{\dots, 0, c_0, c_1, \dots, c_{n-1}, c_n, 0, \dots\}, \quad c_0=c_n=0, \quad (7b')$$

$$y = \{\dots, 0, y_0, y_1, \dots, y_{n-1}, y_n, 0, \dots\}, \quad (7c')$$

and

$$a = \{\dots, 0, a_0, a_1, \dots, a_{n-1}, a_n, 0, \dots\}; \quad (7d')$$

but this particular extension for c, y and a is not dictated by Reinsch's equations (except for the end conditions $c_0=c_n=0$) which entail no assumptions regarding the behavior of y outside the data interval. We use it only for definiteness and convenience. More general extensions will be discussed in section V. The main point here is to represent Reinsch's computations as exactly as possible.

Next notice that while (3) equates two vectors of length n-1, $r_p * c$, where * denotes convolution, and $q * y$ will in general have n + 3 non-zero elements. To represent (3) accurately the extra terms generated by the convolutions must be subtracted from both sides before they are set equal. Thus, (3) becomes

$$r_p * c - c'_p = p (q * y - y') \quad (3')$$

with

$$c_p' = \{ \dots, 0, c_1, [(-4+p/3)c_1 + c_2], 0, \dots, 0, [c_{n-2} + (-4+p/3)c_{n-1}], c_{n-1}, 0, \dots \} \quad (10)$$

and

$$y' = \{ \dots, 0, y_0, (-2y_0 + y_1), 0, \dots, 0, (-2y_{n-1} + y_n), y_n, 0, \dots \}. \quad (11)$$

Note that the prime will not be used to denote differentiation in this paper.

In the case of equation (4) the extra terms are retained. It becomes

$$a = y - \frac{1}{p} q * c, \quad p > 0. \quad (4')$$

Finally, to achieve symmetry in the transforms, and for representational purposes, the discrete vectors will be represented as weighted sums of delta functions. Letting α denote any of them define

$$\alpha(t) = \sum_{i=-\infty}^{\infty} \alpha_i \delta(t-i). \quad (12)$$

The explicit dependence on t will identify continuous time domain functions.

Fourier transforming (3') and (4') modified according to (12) gives

$$r_p(f)c(f) - c_p'(f) = p [q(f)y(f) - y'(f)] \quad (13)$$

and
$$a(f) = y(f) - \frac{1}{p}q(f)c(f), \quad p > 0. \quad (14)$$

Solving (13) for $c(f)$, substituting in (14) and rearranging gives

$$a(f) = u_p(f)y(f) + u_p'(f) [y'(f) - \frac{1}{p}c_p'(f)], \quad p > 0, \quad (15)$$

where
$$u_p(f) = \frac{(p/6)\cos 2\pi f + (p/3)}{\cos^2 2\pi f + ((p/6)-2)\cos 2\pi f + ((p/3)+1)} \quad (16)$$

and
$$u_p'(f) = \frac{2(\cos 2\pi f - 1)}{\cos^2 2\pi f + ((p/6)-2)\cos 2\pi f + ((p/3)+1)} \quad (17)$$

which are the frequency response functions.

The functions $u_p(f)$ and $u_p'(f)$ are shown in figures 1 and 2 over the useful range of p .

III THE ROOTS OF THE CHARACTERISTIC POLYNOMIAL

The characteristic polynomial of the system is the (common) denominator of $u_p(f)$ or $u_p'(f)$ expressed in terms of $z = \exp(2\pi if)$. The roots of this function will be needed in section IV. Setting the denominator of $u_p(f)$ equal to zero and solving for $\cos 2\pi f$ gives

$$z_p^+ = 1 - (p/12) + i(p(72-p))^{1/2}/12 \quad (18)$$

and

$$z_p^- = 1 - (p/12) - i(p(72-p))^{1/2}/12 \quad (19)$$

The principle root is used here and always. These roots lie on the circle of radius 3 centered on the real axis at -2 for $0 \leq p \leq 72$.

For $p \geq 72$ $\text{Im } z_p = 0$. Also, to avoid ambiguity, $\lim z_p^-$ as $p \rightarrow \infty$ is defined to be -2.

If $z = \exp(2\pi if)$ is a root, so is $1/z$ since $\cos 2\pi f = (\exp(2\pi if) + \exp(-2\pi if))/2$. Using this relationship and its converse, the four roots are, for $0 < p < 72$,

$$z_p^{++} = z_p^+ + i(1 - z_p^{+2})^{1/2} \quad (20)$$

$$z_p^{+-} = z_p^+ - i(1 - z_p^{+2})^{1/2} \quad (21)$$

$$z_p^{-+} = z_p^- + i(1 - z_p^{-2})^{1/2} \quad (22)$$

and
$$z_p^{--} = z_p^- - i(1 - z_p^{-2})^{1/2} \quad (23)$$

For $p \geq 72$ there are four distinct negative real roots. Two of these, z_p^{++} and z_p^{-+} , remain within the unit circle,

$$z_p^{++} = z_p^+ + (z_p^{+2} - 1)^{1/2} \quad (24)$$

and

$$z_p^{-+} = z_p^- + (z_p^{-2} - 1)^{1/2} \quad (25)$$

while the reciprocals, z_p^{-+} and z_p^{--} , respectively, remain outside. The $\lim z_p^{++}$ as $p \rightarrow \infty$ is defined to be 0, while $\lim z_p^{-+}$ as $p \rightarrow \infty$ is $-2 + \sqrt{3}$.

The root z_p^{++} is plotted in figure 3. The other roots are simply related to z_p^{++} . Thus z_p^{+-} is the reciprocal, z_p^{-+} is the complex conjugate, and z_p^{--} is the reciprocal of the complex conjugate.

IV THE IMPULSE RESPONSE FUNCTIONS

Using the results of section III, the impulse response functions can be obtained as the inverse Fourier transforms of the frequency response functions $u_p(f)$ and $u_p'(f)$. Decomposing the right side of (16) into partial fractions gives

$$u_p(f) = \frac{A_p}{\cos 2\pi f - z_p^+} + \frac{B_p}{\cos 2\pi f - z_p^-}, \quad (26)$$

where

$$A_p = \frac{(p/6)z_p^+ + (p/3)}{z_p^+ - z_p^-} \text{ and } B_p = \frac{(p/6)z_p^- + (p/3)}{z_p^- - z_p^+}. \quad (27)$$

Then

$$u_p(t) = A_p \int_{-\infty}^{\infty} \frac{\exp(2\pi i f t)}{\cos 2\pi f - z_p^+} df + B_p \int_{-\infty}^{\infty} \frac{\exp(2\pi i f t)}{\cos 2\pi f - z_p^-} df, \quad (28)$$

$$= \frac{A_p}{\pi i} \int_C \frac{z^t dz}{z^2 - 2z_p^+ z + 1} + \frac{B_p}{\pi i} \int_C \frac{z^t dz}{z^2 - 2z_p^- z + 1} \sum_{i=-\infty}^{\infty} \delta(t-i). \quad (29)$$

the unit circle, C , being traversed in the positive sense for $t > 0$. To see that (29) follows from (28) consider (12) and the fact that the series of delta functions used to repeat the principal part of $u_p(f)$ by convolution is the transform of the series in t .

The integrals in (29) can be evaluated by the method of residues [1]. The roots of the denominator of the first integrand are z_p^{++} and z_p^{+-} , and for the second z_p^{-+} and z_p^{--} , which are respectively complex conjugates of the first (reciprocal) pair. Then, for all $p > 0$, the impulse response function

$$u_p(t) = (2K_p z_p^{++|t|} + 2L_p z_p^{-+|t|}) \sum_{i=-\infty}^{\infty} \delta(t-i) = v_p(t) \sum_{i=-\infty}^{\infty} \delta(t-i), \quad (30)$$

where

$$K_p = A_p / (z_p^{++} - 1/z_p^{++}), \quad (31)$$

and

$$L_p = B_p / (z_p^{-+} - 1/z_p^{-+}). \quad (32)$$

Using the same conjugacy relationships between the roots, for $0 < p < 72$,

$$v_p(t) = 4\text{Re}\{K_p z_p^{++|t|}\} = 4|K_p| |z_p^{++}|^{|t|} \cos(\text{Arg } K_p + |t| \text{Arg } z_p^{++}) \quad (33)$$

In the limit as $p \rightarrow 0$, $v_p(t) \rightarrow 0$ uniformly everywhere.

For $p > 72$ there are four distinct negative roots occurring in reciprocal pairs. The impulse response function becomes the sum of two complex numbers with different amplitudes (decaying exponentially in $|t|$) rotating in opposite senses around the t axis, piercing the imaginary- t plane at half integral values of t and the real- t plane at integral values of t . In the limit as $p \rightarrow \infty$, $v_p(t) \rightarrow 1$ for $t=0$ and $\rightarrow 0$ elsewhere. That is, $u_p(t)$ is a spike at the origin.

The function $v_{72}(t)$ must be evaluated by a limit process. The limits from the right and the left are the same and thus $v_p(t)$ is continuous at $p=72$ for all values of t . The function $v_p(0)$ in particular is plotted in figure 4.

The functions $v_p(t)$ for $\log p = -3, -4, -5$ are plotted in figure 5.

The analysis for u_p' proceeds as for u_p except that the constants A_p' and B_p' are different.

$$A_p' = (2z_p^+ - 2)/(z_p^+ - z_p^-) \quad (34)$$

and

$$B_p' = (2z_p^- - 2)/(z_p^- - z_p^+) \quad (35)$$

$v_p'(0)$ is plotted in figure 4, and the functions $v_p'(t)$ for $\log p = -3, -4, -5$ are plotted in figure 6.

V DISCUSSION AND AN EXAMPLE

Transforming (15) back into the time domain and reverting to the discrete vector symbolism gives

$$a = u * y - u_p' * (y' - \frac{1}{p} c_p'), \quad p > 0, \quad (36)$$

and the operation of the smoothing spline is seen to be representable in the time domain as the sum of two convolutions. The first term accurately represents the operation when end effects can be ignored; for instance far enough away from the ends of the data so that the filter cannot feel them, or in the cases in which a reasonable extrapolation of the data can be made. The second term is peculiar to the smoothing spline and accounts for its apparent lack of end effects. It is necessary to consider the sense in which the smoothing spline has no end effects.

This phrase is appropriate if one accepts the global variational definition of the smoothing spline as implying that its quasi-local behavior is the same at the ends of the data set as near the center.

Indeed the Euler-Lagrange equations must be satisfied at all interior points. In this sense the second term in (36) "fixes" the computations so that u_p alone is an adequate description of its operation, in which case one need only consult figure 1 in selecting an appropriate p value. We have been using the spline in this sense. It is more appropriate in this paper, however, to consider the end behavior in terms of the convolution representation.

If a different choice had been made in extending the vectors c , y and a , the contribution of the second term in (36) would have been different. In fact there are (an infinite number of) extensions such that this second term would contribute nothing. They can be computed by convolving u_p with y and requiring that the resulting values be equal to a_i for $0 < i < n$. If one accepts these extensions as "reasonable", then again u_p by itself adequately describes the filter. This implies that the smoothing spline can be used as a predictor or to achieve higher resolution in spectral estimates. If the physical model entails that the data vary according to some principle of minimum curvature at least at lower frequencies such an idea is certainly tenable, but as in other kinds of prediction and spectral enhancement (maximum entropy, for instance) due consideration must be given to the underlying model. Prediction is reliable only if the model is appropriate.

Four smoothings of tree-ring width sequences are shown in figure 7. The same figure appears in [2]. The dashed lines show orthogonal polynomial fits generated by the INDXA program [5] the orders of which are determined by stopping when an additional order fails to reduce the residual variance by 5% or more. The log p value was -4 in each case.

Tree-ring width sequences from forest interior sites tend to exhibit behavior similar to that seen in records of complex physical processes such as level changes in streams and reservoirs and climatic variations. These data exhibit sudden changes in level (episodic behavior) over a wide range of time scales. The spectra tend to be lumpy and have concentrations of power near zero frequency. The particular suitability of the smoothing spline as a low pass filter for certain tree-ring records, and perhaps for other similar data, is accounted for by the shape of the impulse response functions. The spikey exponential envelope with the $|t|$ dependence ensures that the filter will be responsive to sudden changes in level. On the other hand, the fact that it is a low-pass filter with a minimum curvature property ensures that power near zero frequency will be preserved. The frequency response function is smooth at the origin. Thus trends and slow wandering in the data will be preserved. Selecting a value of S (or p) affects a compromise between fitting and smoothing the data. We see how this compromise manifests itself in the digital filters.

Programming of equations (3') and (4') can be done in various ways. We have not tried to optimize the computations beyond using an equation solver specialized for symmetric band matrices. The user concerned about storage might try a Toeplitz-specific algorithm. The only caution regards the possible necessity of using double precision accumulation variables. We found that in single precision (32 binary bits) the computations were unstable for values of p smaller than 10^{-4} roughly.

APPENDIX A

It is not obvious that the weight function shown in figure 6 will, in the limit of small p , give a series of values which lie on a straight line for $0 \leq i \leq n$ and be zero outside this interval when convolved with anything at all. The purpose of this appendix is to examine this limiting case.

For $0 < p < 72$, and thus in the limit as $p \rightarrow 0$,

$$v_p'(t) = 4 \operatorname{Re}\{K_p' z_p^{++|t|}\} = 4 |K_p'| |z_p^{++}|^{|t|} \cos(\operatorname{Arg} K_p' + |t| \operatorname{Arg} z_p^{++}). \quad (A1)$$

Letting r , θ' and θ denote small positive numbers, for small p A1 can be written

$$v_p'(t) = 4 |K_p'| (1-r)^{|t|} \cos((3\pi/4) - \theta' - |t|\theta). \quad (A2)$$

Expanding in r and θ' and θ ,

$$v_p'(t) = k [1 - |t|r + o(r^2)] [1 - \theta' - |t|\theta - o(\theta^2)], \quad (A3)$$

where

$$k = -4 |K_p'| / \sqrt{2}.$$

To first order therefore

$$v_p'(t) \doteq k(1 - \theta' - s|t|), \quad \text{for } s \ll |t| \ll 1/s, \quad (A4)$$

where $s = r + \theta$.

Denoting the four non-zero values of $(y' - \frac{1}{p} c_p')$ by w_{-1} , w_0 , w_n , w_{n+1} , at least the following four requirements must be met:

$$k(1 - \theta')w_{-1} + k(1 - \theta' - s)w_0 = 0, \quad (A5)$$

$$k(1 - \theta' - s)w_{-1} + k(1 - \theta')w_0 = a_0, \quad (A6)$$

$$k(1 - \theta')w_n + k(1 - \theta' - s)w_{n+1} = a_n, \quad (A7)$$

$$\text{and } k(1 - \theta' - s)w_n + k(1 - \theta')w_{n+1} = 0. \quad (A8)$$

Since in the limit of small p , $|K_p'|$ and thus $|k|$ become very large, k can be divided out of A5 - A8. Solving the resulting equations for w_{-1} and w_0 (A5 and A6) and w_n and w_{n+1} (A7 and A8) yields to first order in s ,

$$w_{-1} \doteq -a_0(1-\theta'-s)/2s(1-\theta'), \quad (\text{A9})$$

$$w_0 \doteq a_0/2s, \quad (\text{A10})$$

$$w_n \doteq a_n/2s, \quad (\text{A11})$$

and $w_{n+1} \doteq -a_n(1-\theta'-s)/2s(1-\theta'). \quad (\text{A12})$

Since $s > 0$ for all finite p these solutions are defined for the limit.

Also, by the symmetric linear (to first order) dependence of $v_p'(t)$ on t , it follows from A5 that $a_i \doteq 0$ for all $-1/s \ll i < -1$; from A8, that $a_i \doteq 0$ for all $n+1 < i \ll 1/s$; and from A6 and A7 that a_i varies linearly with i (to first order) for $0 \leq i \leq n$. In the limit as $p \rightarrow 0$ the first order expressions become exact and $1/s \rightarrow \infty$.

APPENDIX B

The first terms in the central finite difference expansions of the second and fourth derivative operators are obvious in the matrices Q and R. Corresponding to equations (3') and (4') there are the coupled differential equations

$$[D^4 + (p/3)(D^2 + 6)] C(t) = pD^2 Y(t) \quad (B1)$$

and

$$A(t) = y - \frac{1}{p} D^2 C(t), \quad p > 0. \quad (B2)$$

Conversely, a finite difference solution of (B1) and (B2) will satisfy equations (3') and (4'). More generally, the theory of Reinsch's smoothing splines is related to the theory of the interpolated finite difference solutions of certain ordinary differential equations with constant coefficients. The remainder of this appendix is a brief analysis of the equations (B1) and (B2), analogous to that done for the spline equations, which will facilitate some comparisons the reader may wish to make between the continuous and discrete systems.

Fourier transforming equations (B1) and (B2) and solving for A(f) in terms of Y(f), gives

$$A(f) = \frac{p(3-2\pi^2 f^2)}{24\pi^4 f^4 + p(3-2\pi^2 f^2)} Y(f) = U_p(f) Y(f). \quad (B3)$$

Note that this same transform would arise from the single equation

$$(D^4 + (p/3)(D^2 + 6))A(t) = (p/3)(D^2 + 6)Y(t). \quad (B4)$$

This won't work in the discrete case because of the correction terms.

The roots of the denominator for $0 \leq p \leq 72$ are

$$f_p^{2+} = (p + i(p(72-p))^{1/2})/24\pi^2 \quad (B5)$$

and

$$f_p^{2-} = (p - i(p(72-p))^{1/2})/24\pi^2 \quad (B6)$$

and therefore

$$f_p^{++} = +(f_p^{2+})^{1/2}, \quad (B7)$$

$$f_p^{+-} = -(f_p^{2+})^{1/2}, \quad (B8)$$

$$f_p^{-+} = +(f_p^{2-})^{1/2}, \quad (B9)$$

and

$$f_p^{--} = -(f_p^{2-})^{1/2}. \quad (B10)$$

Corresponding to the constants A_p and B_p in the inverse transform there are

$$\alpha_p = p(3-2\pi^2 f_p^{2+})/24\pi^4 (f_p^{2+} - f_p^{2-}), \quad (B11)$$

and

$$\beta_p = p(3-2\pi^2 f_p^{2-})/24\pi^4 (f_p^{2-} - f_p^{2+}), \quad (B12)$$

while

$$k_p = \alpha_p / f_p^{++}, \quad (B13)$$

and

$$l_p = \beta_p / f_p^{-+}. \quad (B14)$$

The contour of integration consists of the real axis and the upper half of the point at infinity, traversed counterclockwise for $t > 0$.

For $0 < p < 72$

$$U_p(t) = -2\pi \exp(-2\pi \text{Im}\{f_p^{++}\} |t|) \text{Im}\{k_p \exp(2\pi i \text{Re}\{f_p^{++}\} |t|)\}. \quad (B15)$$

For $p \geq 72$, each contour ($t > 0$) includes two of the real roots and

$$U_p(t) = -\pi(k_p \sin 2\pi f_p^{++} |t| + l_p \sin 2\pi f_p^{-+} |t|). \quad (B16)$$

There is no decay in the continuous case for $p \geq 72$.

Evaluation of the limits as p approaches 72 from the left and right will show that for any value of t , $U_p(t)$ is continuous at $p=72$.

ACKNOWLEDGEMENTS

The authors wish to thank Paul Stoffa, John Shepherd and Keith McCamy for many constructive criticisms. Funds for this work were provided through the Climate Dynamics Program of the National Science Foundation, grant no. ATM-77-19217.

FIGURE CAPTIONS

- Figure 1. The gain of the frequency response function $u_p(f)$ for various values of the parameter p over the range $.001 \leq f \leq .5$.
- Figure 2. The base ten logarithm of the negative of the frequency response function $u_p'(f)$ for various values of the parameter p over the range $.001 \leq f \leq .5$. The dashed line passes through the maximum (circles) on these curves.
- Figure 3. The root z_p^{++} of the characteristic polynomial $r_p(z)$. Values of the parameter p are shown along the curve.
- Figure 4. The central values $v_p(0)$ and $v_p'(0)$ of the smooth parts of the impulse response functions u_p and u_p' .
- Figure 5. The smooth part $v_p(t)$ of the impulse response function u_p for three values of p most useful for standardizing forest interior tree-ring series.
- Figure 6. The smooth part $v_p'(t)$ of the impulse response function u_p' for three values of p most useful for standardizing forest interior tree-ring series.
- Figure 7. Four examples of the smoothing spline applied to forest interior ring-width series. Each spline was computed for $\log p = -4.0$. The solid curve is the spline fit and the dashed curve is the orthogonal polynomial fit from program INDXA (see ref. 4). The four series are from the same site in southeastern New York and the lower two plots are curves from the same tree.

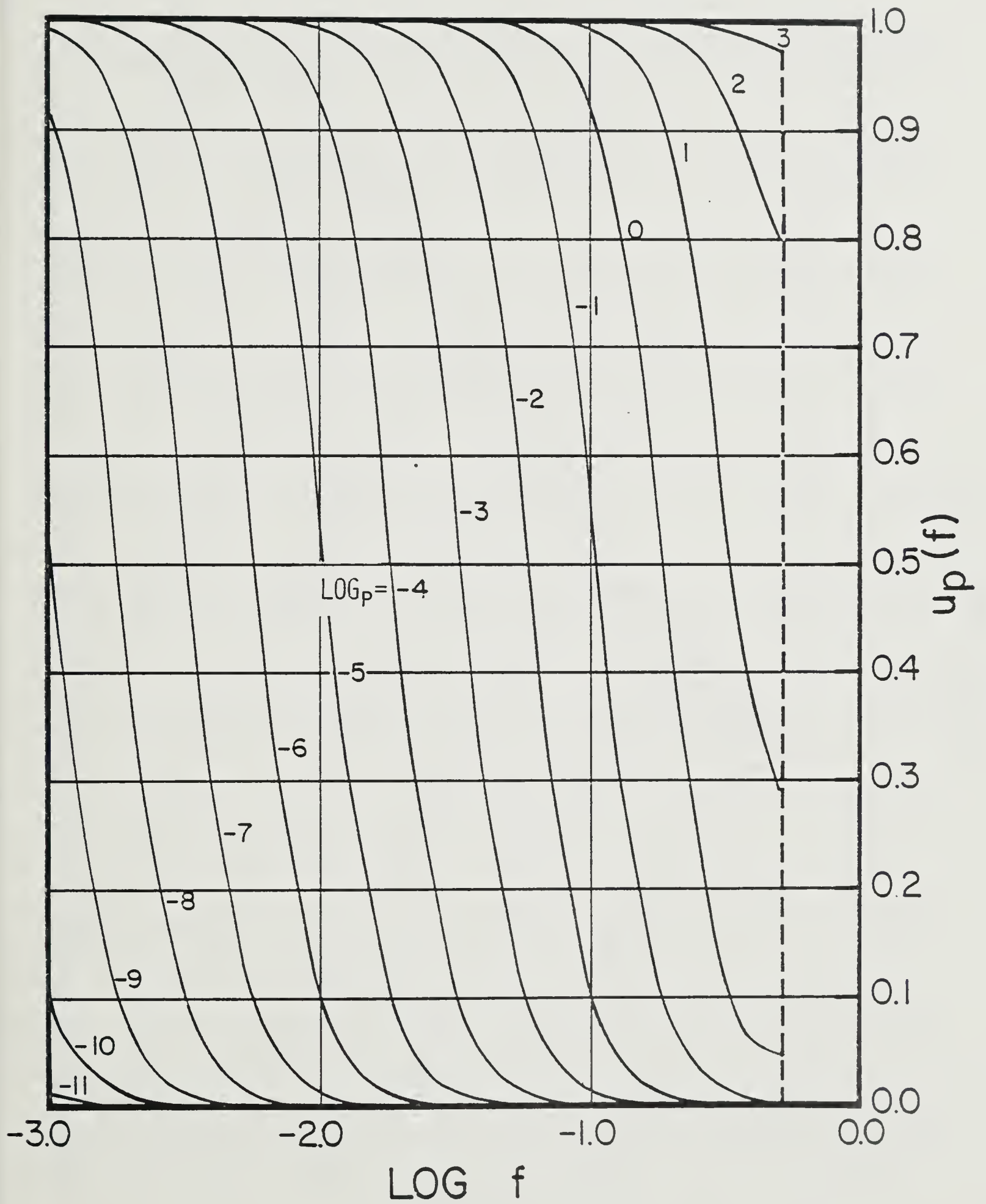


Figure 1

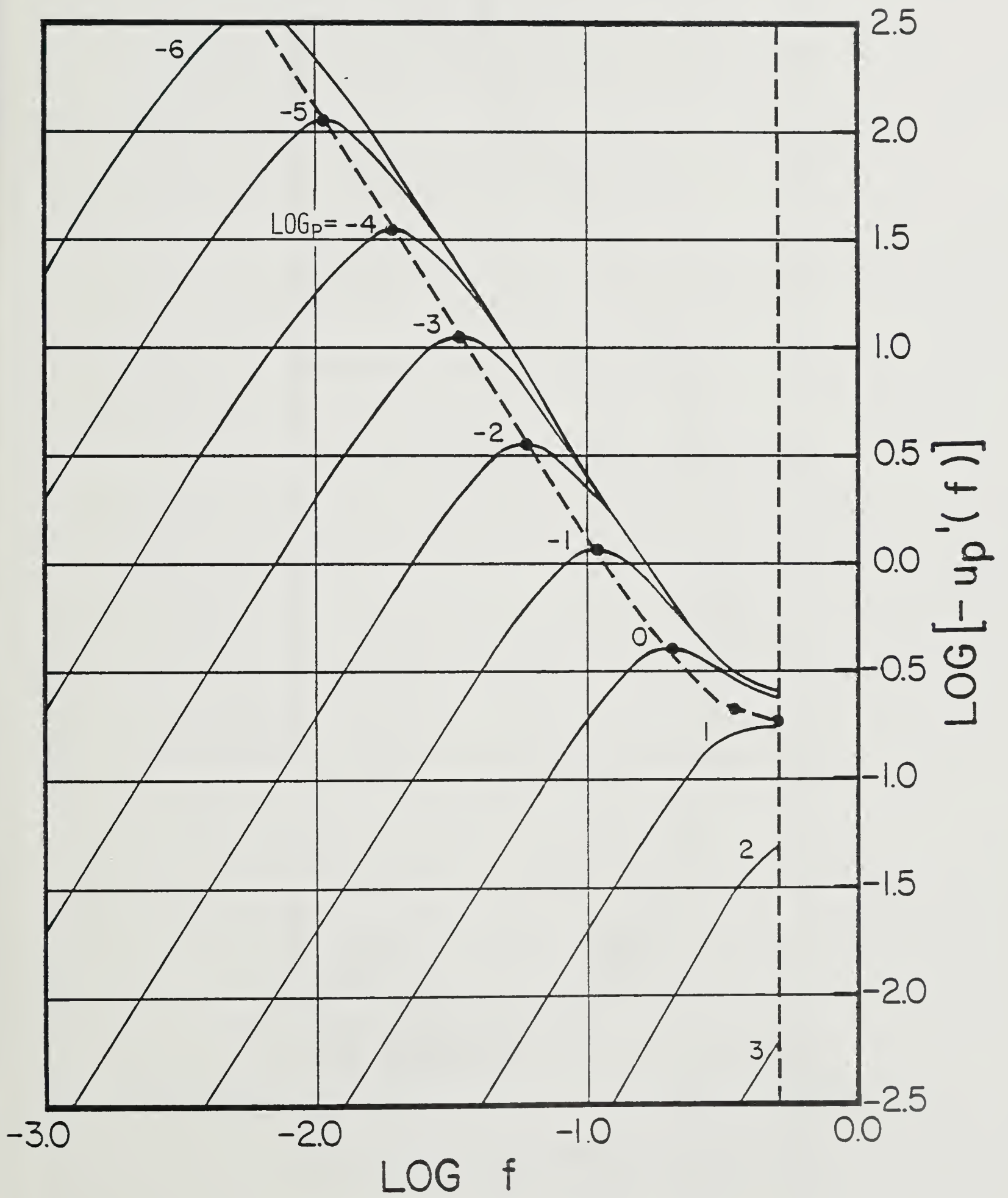


Figure 2

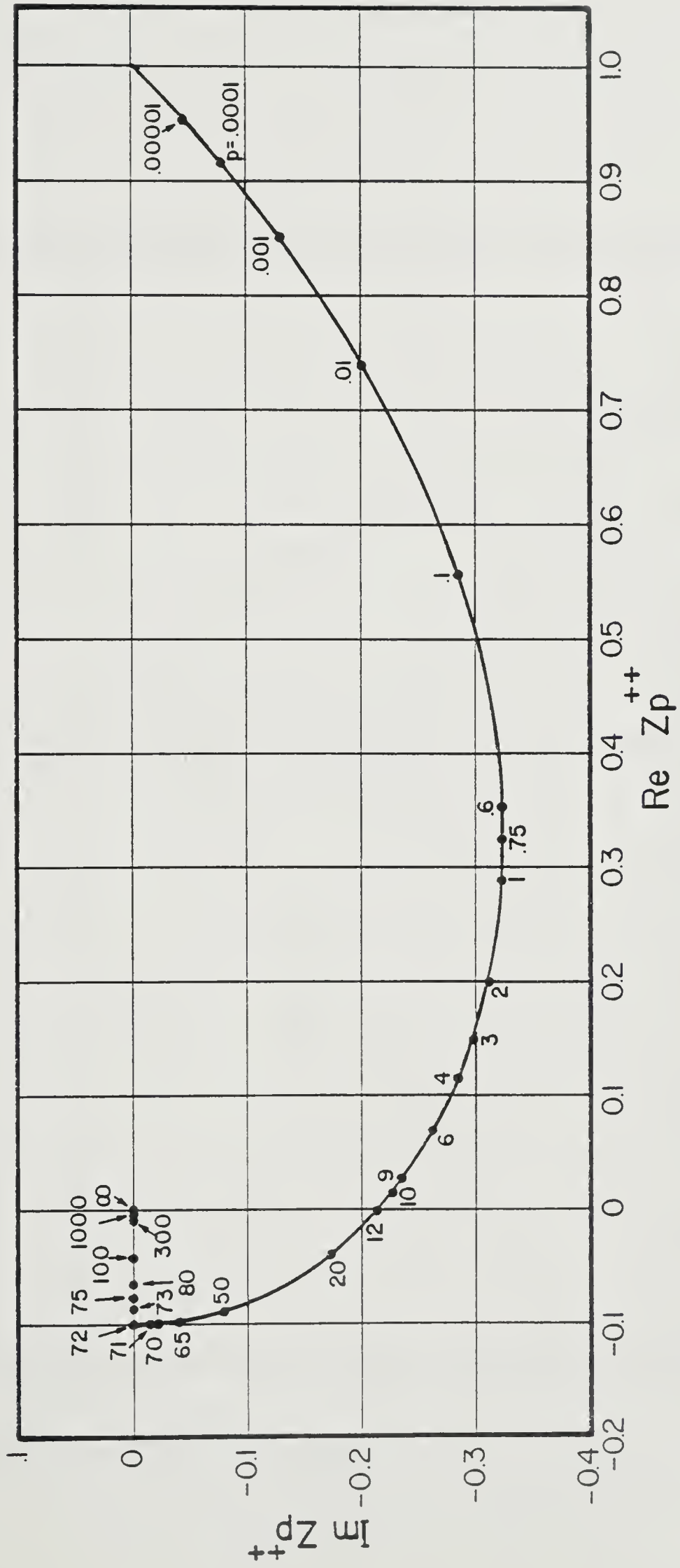
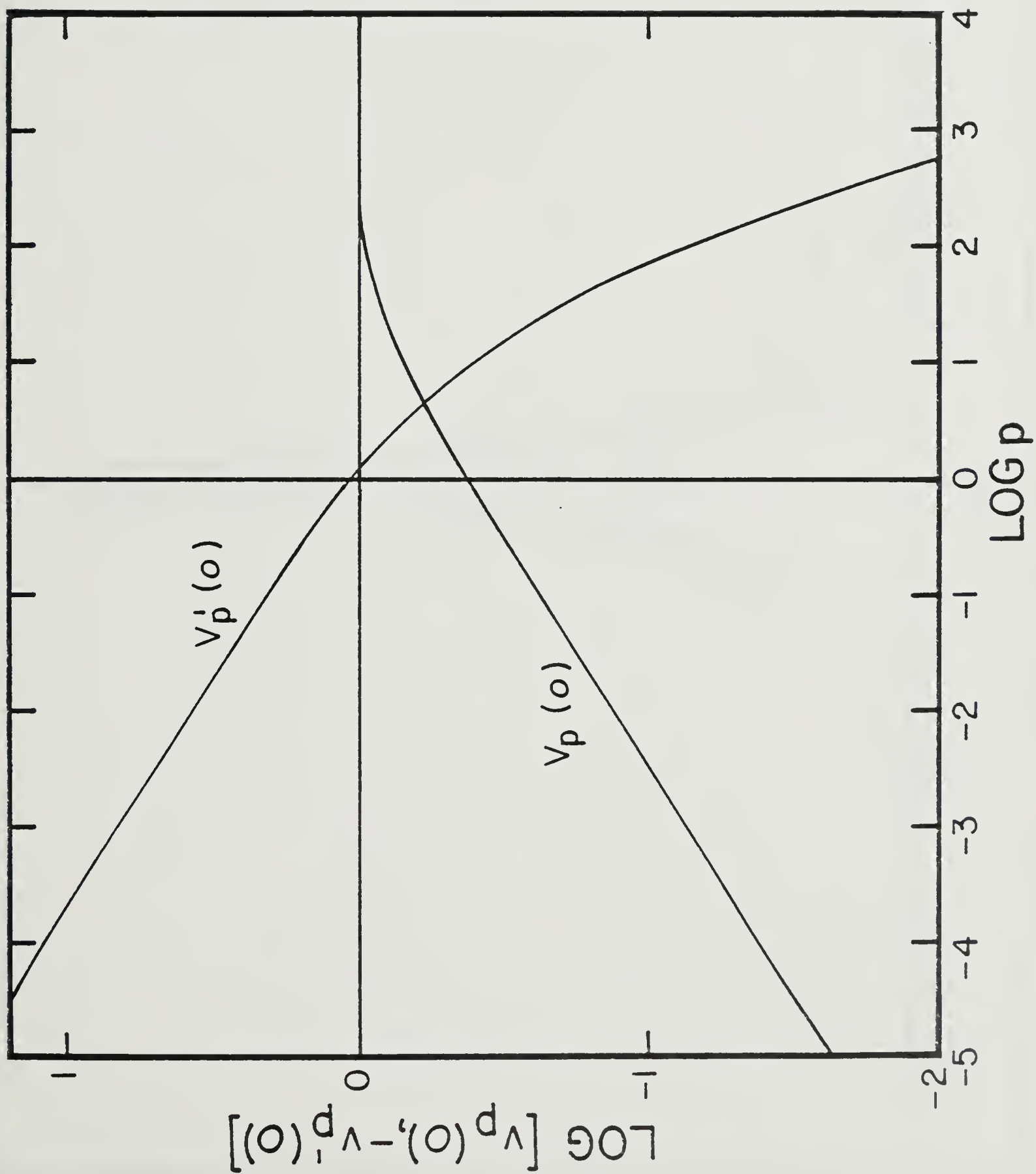
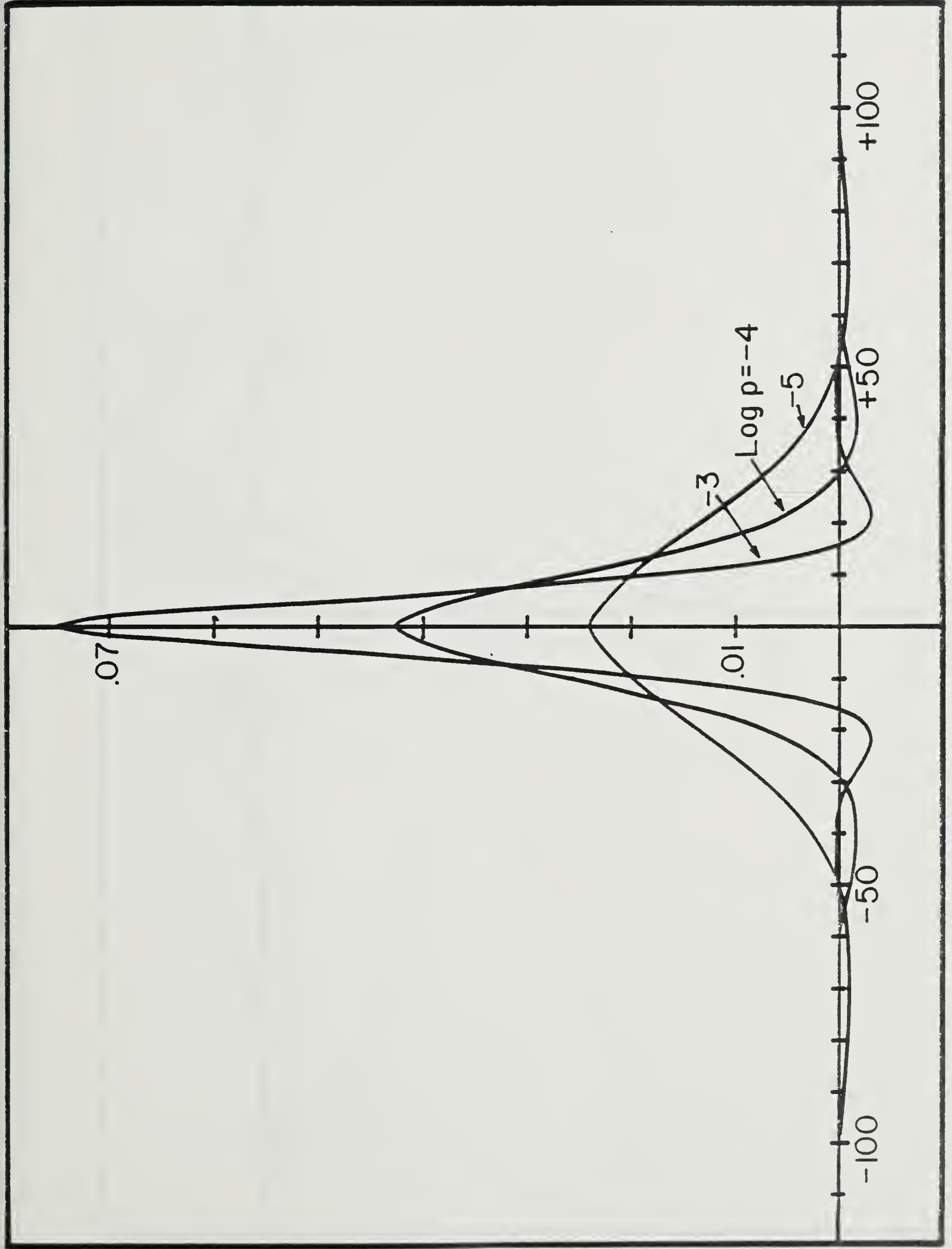


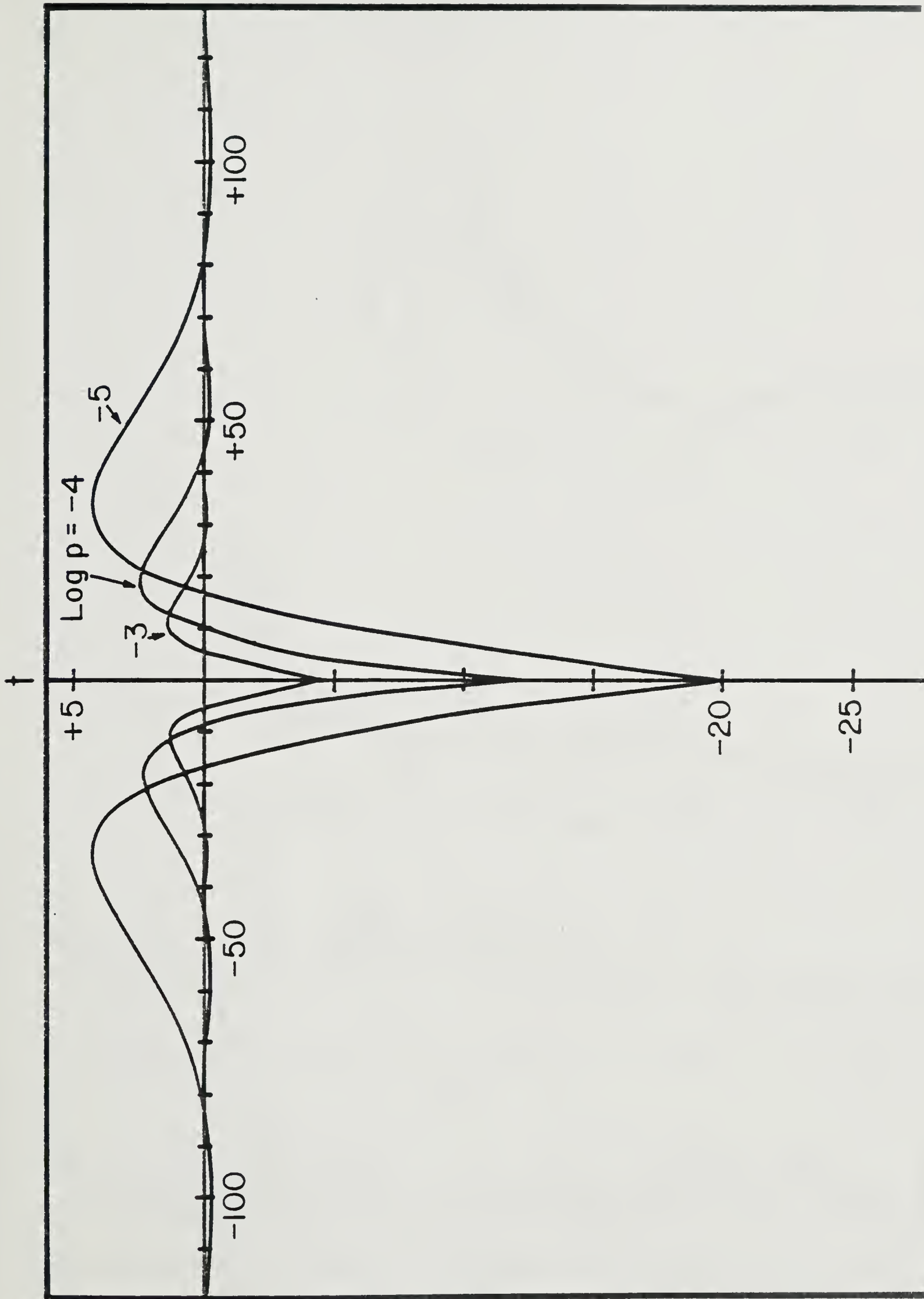
Figure 3





$v_d(t)$

$v_p(t)$



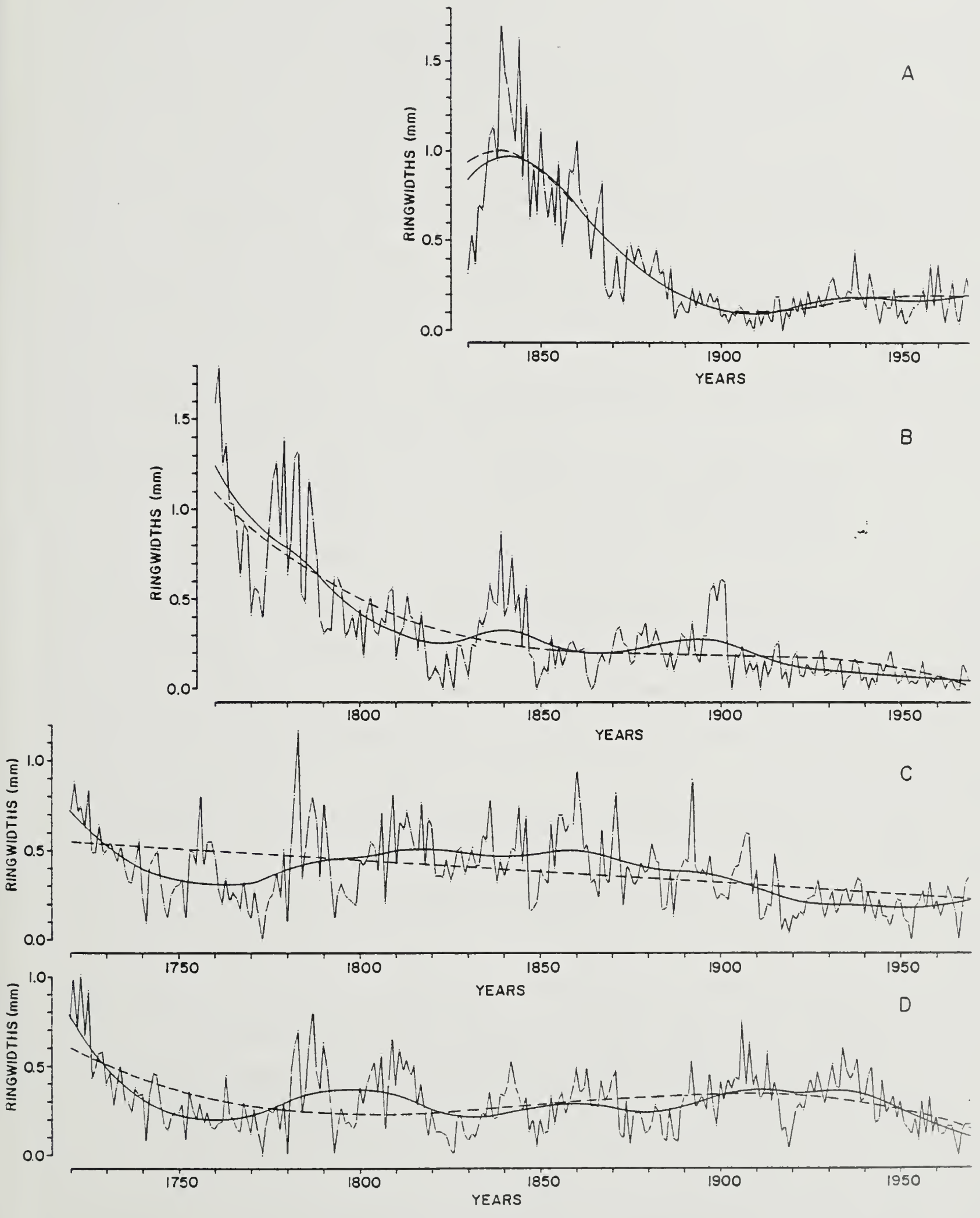


Figure 7

REFERENCES

1. Churchill, R.V.: Introduction to complex variables and applications, New York: McGraw-Hill Book Company, Inc., 1948.
2. Cook, E.R. and Peters, K.: The smoothing spline: a new approach to standardizing forest interior ring-width series for dendroclimatic studies. (submitted).
3. Craven, P.G.: Smoothing noisy data with spline functions, Numer. Math. 31, 377-403 (1979).
4. Demmler, A., Reinsch, C.: Oscillation matrices with spline smoothing, Numer. Math. 24, 375-382 (1975).
5. Fritts, H.C., Mosimann, J.E. and Bottorff, C.P.: A revised computer program for standardizing tree-ring series. Tree-Ring Bulletin 29, 15-20 (1969).
6. Horowitz, L.L.: The effects of spline interpolation on power spectral density. IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. AAP-22, number 1, 22-26 (February 1974).
7. Reinsch, C.H.: Smoothing by spline functions, Numer. Math. 10, 177-183 (1967).
8. Reinsch, C.H.: Smoothing by spline functions, II. Numer. Math. 16, 451-454 (1971).
9. Wahba, G.: Smoothing noisy data with spline functions, Numer. Math. 24, 383-393 (1975).
10. Wahba, G., Wold, S.: A completely automatic French curve: Fitting spline functions by cross-validation, Comm. Stat. 4, 1-17 (1975).
11. Wahba, G., Wold, S.: Periodic splines for spectral density estimation: The use of cross-validation for determining the degree of smoothing, Comm. Stat. 4, 125-141 (1975).
12. Wold, S.: Spline functions in data analysis, Technometrics 16, 1-11 (1974).

LAMONT-DOHERTY GEOLOGICAL OBSERVATORY
OF COLUMBIA UNIVERSITY

