Ideal Wave Filters *

By H. W. BODE and R. L. DIETZOLD

The increasing usefulness of wave filters in the telephone plant, together with rising standards of quality, emphasizes the need of a systematic method for approximating ideal characteristics as closely as we please. By an ideal filter is meant a network having the properties of a distortionless transducer over a given frequency range and suppressing all other frequencies. A design method is presented whereby an arbitrarily close approximation to these properties may be realized in a physical network. Examples of actual designs illustrate the engineering features involved in the practical application of the theory.

Introduction

In the phenomenal advance of telephone practice during the past twenty years, almost every step has further restricted the distortion which individual parts of a transmission system can be allowed to introduce into the signal. The extension of circuits to great distances made it necessary that each link pass on to the next a more faithful copy of the signal so that the accumulated effects of many links might not endanger the intelligibility. The extension of telephone circuits to new uses, such as the transmission of pictures and the distribution of broadcast programs, imposed new demands for accuracy. Each of these has required rising standards of performance for wave filters. More than anything else, however, it has been the introduction of carrier methods, with their comparatively large utilization of selective structures, which has given prominence to the problem of reducing the distortion from wave filters. With the increase in length and complexity of carrier systems, the problem of providing wave filters which will have no harmful effect upon transmission has become one of increasing importance.

What this requires of the filters quickly appears if we recall that a structure which transmits *all* signals without distortion must (1) possess a characteristic impedance which is a pure resistance independent of frequency; (2) attenuate steady sinusoidal signals equally at all values of frequency; and (3) introduce a rotation in phase proportional to the frequency. In filter theory we need consider these requirements over only a limited band, since the signals which filters

* The reader is referred to the preceding paper entitled "A General Theory of Electric Wave Filters."

are meant to transmit, whether voice, telegraph, or television, are of a specified type having energies concentrated in certain portions of the frequency spectrum. We can therefore say that an ideal filter is one which has the ideal phase, impedance and attenuation properties in the frequency range of the desired signal and which totally suppresses all other frequencies.

The conventional ladder type filter structures which have been so extensively studied may be made to yield any desired suppression at the unwanted frequencies. In the range of wanted frequencies, however, they show wide departures from all three ideal properties. The impedance characteristic can be greatly improved by suitable elaboration of the filter structure itself, but to approximate uniformity of loss or linearity of phase shift it has been necessary to make use of supplementary networks of empirical design.[1]

The design of such corrective networks is by no means an easy task, primarily because the filter characteristics for which they are supposed to compensate change very rapidly with frequency in certain intervals. Nevertheless, much has been achieved. Thus it has been found possible to limit reflection coefficients to 2 per cent, in contrast with coefficients of 50 per cent not uncommonly tolerated in the systems of ten years ago. Improvements in the other characteristics have been comparable. In modern systems variations in attenuation of a few hundredths of a decibel, or in phase slope of a few per cent, can be attained if need be. These limits, however, demand the most patient and skillful design, and can seldom be met unless control of a single one of the characteristics is especially important. Since amplitude equalizers introduce non-linear phase, phase correctors non-uniform loss, and so on, the problem becomes increasingly difficult when close requirements must be met in several characteristics simultaneously.

By contrast, the method proposed in this paper gives the various characteristics simultaneously in a single network without recourse to auxiliary corrective structures. The method is a systematic one, requiring comparatively little in the way of cut and try design work. At the same time it preserves a measure of the flexibility of the existing technique, so that when considerable deviation from the ideal is tolerable in one or more characteristics, a corresponding economy of materials may be effected.

[1] The distortion problem has been discussed by several writers in this *Journal*. See, for example, S. P. Mead, "Phase Distortion and Phase Distortion Correction," April, 1928, p. 195; O. J. Zobel, "Distortion Correction in Electrical Circuits . . .," July, 1928, p. 438; C. E. Lane, "Phase Distortion in Telephone Apparatus," July, 1930, p. 493; E. B. Payne, "Impedance Correction of Wave Filters," October, 1930, p. 770.

The discussion which follows has a two-fold objective. The first is purely theoretical: to demonstrate that no matter how close the limits of deviation from the ideal may be set, there is a finite physical network all of whose characteristics meet these limits, except within a certain "transition interval" about each cut-off, which transition interval may also be taken as narrow as we please. This is by no means trivial; for it is known that no network, finite or infinite, can meet the ideal characteristics *exactly*.[2]

The second object is to guide the selection, from among the many networks which would meet the requirements of a given practical problem, of that one which meets them most economically. This part of the paper contains a number of examples, among them some which illustrate the use of slight empirical variations as a means of obtaining the highest measure of economy when wide deviations from the ideal are more tolerable in one respect than in others. The final example, which is segregated as Part III, deals with a situation met in picture transmission circuits, where the selectivity required is frequently small, but the effects of phase distortion may be very serious. Here a modification of the design technique leads to a filter which has comparatively modest selectivity but which exhibits a linear phase characteristic not only in the transmitting band but also in the range of rising attenuation.

Part I—Theoretical Analysis

Since linear phase shift is not available from ladder networks, the analysis will be based upon the more flexible lattice configuration. Although the lattice lends itself particularly well to the theoretical design problem, it is not so satisfactory for purposes of physical construction. After the paper design has been made, therefore, it will usually be desirable to convert it to a more suitable practical configuration. This can be done by methods described elsewhere.[3]

We may greatly simplify the theoretical discussion by ignoring the effects of parasitic dissipation—a simplification warranted by Mayer's Theorem,[4] which states that the attenuation resulting from dissipation

[2] This proposition is due to Dr. T. C. Fry, who showed that in a transducer possessing the steady-state characteristics of an ideal filter, a signal would arrive at the receiving terminals before it began to be impressed on the sending terminals. As this is absurd, we must conclude that no such system exists.

[3] H. W. Bode, "A General Theory of Electric Wave Filters," M.I.T. *Journal of Mathematics and Physics*, November, 1934. A summary of this article appears in this issue of the *Bell System Technical Journal*.

[4] H. F. Mayer, "Über die Dämpfung von Siebketten im Durchlässigkeitsbereich," *E. N. T.*, October, 1925, p. 335. His results were later somewhat extended by Feige and Holtzapfel, "Dämpfung und Winkelmass von Vierpolen mit geringen Verlusten," *T. F. T.*, July, 1932, p. 179. Even these latter results are capable of considerable generalization, so as to include other characteristics of the network besides the transfer constant.

is proportional to the derivative of the phase characteristic. The realization of a linear phase shift in the transmission band therefore automatically carries with it the satisfaction of the requirement of uniform loss in this range.[5] It can also be shown that the other characteristics of the network will not be appreciably affected by slight uniform dissipation.

Moreover, it is well known that the image impedance and transfer constant of a lattice structure are controlled by independent parameters.[6] We can, therefore, dissociate the problem of providing the required constant image impedance in the transmission band from that of providing the required loss and phase characteristics.[7] For the moment we shall fix our attention on the transfer constant.

With these simplifications our problem reduces to that of constructing a filter whose transfer constant on a non-dissipative basis represents a linear phase shift in the transmission band and an infinite loss in the attenuation bands, these being separated by narrow "transition intervals" in the neighborhood of the cut-offs. These transition intervals may be taken small at pleasure, but must be assigned in advance to insure the physical realizability of the network.

Formulation of Requirements—Low-Pass Filters [8]

If the impedances of the arms of a lattice are Z_x and Z_y, Fig. 1, it is well known that the image transfer constant and the image impedance are given by the expressions [9]

$$\tanh\frac{\theta}{2} = \sqrt{\frac{Z_x}{Z_y}}, \tag{1}$$

$$Z_I = \sqrt{Z_x Z_y}. \tag{2}$$

[5] Strictly speaking, a slight qualification should be placed upon this statement. Our process of approximating the ideal characteristics will lead to a phase shift which ripples about the desired linear characteristic, the number of ripples depending upon the number of elements used. As the number of elements is increased indefinitely, the linear characteristic is approximated more and more closely, but it is evidently not a necessary consequence of this that the slope of the ripples should approach constancy. We shall be able to show, however, that with the actual process used, the amplitude of the ripples decreases so rapidly that $dB/d\omega$ approaches constancy as B approaches linearity.

[6] This follows at once from equations (4) and (5), p. 220.

[7] A method of choosing the lattice parameters to give a substantially constant impedance in the transmission band has in fact already been obtained by W. Cauer, "Siebschaltungen," V. D. I. Verlag, Berlin, 1931; or "Ein Interpolationsproblem mit Funktionen mit Positivem Realteil," *Math. Zeit.*, November, 1933, p. 1. An alternative method will eventually be developed as a by-product of the present analysis.

[8] The extension to filters of other types is given on p. 225.

[9] G. A. Campbell, "Physical Theory of the Electric Wave Filter," this *Journal*, Vol. I, No. 2, November, 1922, p. 1.

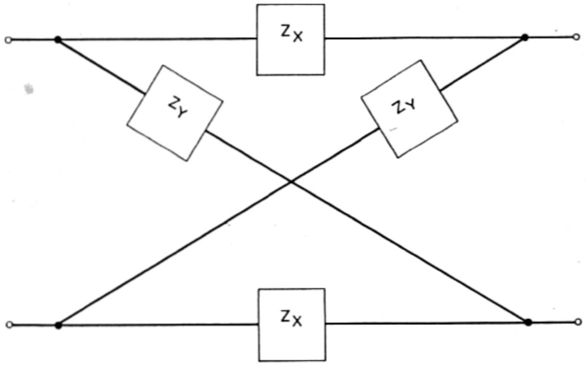


Fig. 1— The symmetrical lattice.

The relation (1) requires for transmission, i.e., for θ imaginary, that Z_x and Z_y differ in sign; for attenuation, i.e., for θ real, that Z_x and Z_y be alike in sign. In the case of the low-pass filter, this amounts to requiring correspondence of zeros (resonances) in one arm to poles (anti-resonances) in the other for $f < f_c$, and of zeros to zeros and poles to poles for $f > f_c$, where f_c, the cut-off, is a critical frequency which appears in one arm only.[10] If we denote these critical frequencies by $f_1, f_2, \cdots, f_r$ in the range below f_c, and by $f_1', f_2', \cdots, f_s'$ in the range above f_c, and if we make use of a well-known theorem [11] we readily find that Z_x and Z_y have forms similar to [12]

$$\left.\begin{aligned} Z_x &= iK_x f\,\frac{a_2a_4\cdots a_{r-1}a_ca_2'\cdots a'_{s-1}}{a_1a_3\cdots a_ra_1'\cdots a_s'}, \\ Z_y &= -\frac{iK_y}{f}\,\frac{a_1a_3\cdots a_ra_2'\cdots a'_{s-1}}{a_2a_4\cdots a_{r-1}a_1'\cdots a_s'}, \end{aligned}\right\} \qquad (3)$$

[10] In the basic theory given by Dr. Campbell, in the paper just referred to, it is shown that in general a lattice having many natural frequencies is a milti-band-pass filter. The extension of the theory in the manner shown above, in which separate parameters for the control of the transfer constant and image impedance are obtained by imposing special conditions on the natural frequencies, thus rendering many bands confluent, was discovered and exploited independently by W. Cauer and one of the present writers (see Cauer, "Siebschaltungen" and later papers; or H. W. Bode, U. S. Patent No. 1828454, also "A General Theory of Electric Wave Filters," loc. cit.). The published work by Dr. Cauer gives a particularly complete discussion. It appears from a recent informal communication from Dr. Campbell to the authors, however, that this extension was also considered by him and was described briefly in the Yale-Harvard Lectures on Wave Filters delivered in 1923. The lectures have unfortunately not been published. Their content, however, is similar to that given in the discussion above.

[11] R. M. Foster, "A Reactance Theorem," this *Journal*, Vol. III, No. 2, April, 1924, p. 259.

[12] The cut-off factor a_c may appear in either the numerator or denominator of either form, and the line- and cross-arms may be interchanged. Otherwise the expression is general.

and hence that [13]

$$\tanh\frac{\theta}{2} = i\sqrt{\frac{K_x}{K_y}}f\frac{a_2a_4\cdots a_{r-1}}{a_1a_3\cdots a_r}\sqrt{a_c}, \tag{4}$$

$$Z_I = \sqrt{K_xK_y}\sqrt{a_c}\frac{a_2'\cdots a'_{s-1}}{a_1'\cdots a_s'}, \tag{5}$$

where the a's are shorthand notation for

$$a_j = 1 - \frac{f^2}{f_j^2}, \qquad a_j' = 1 - \frac{f^2}{f_j'^2}. \tag{6}$$

We shall have frequent occasion to distinguish between those critical frequencies of (4) which lie in the practical transmission band and those which lie in the transition interval. To this end we shall denote that interval by (f_A, f_B), where obviously $f_A < f_c < f_B$, and shall write P for the group of factors

$$P = i\sqrt{\frac{K_x}{K_y}}f\frac{a_2a_4\cdots a_{A-1}}{a_1a_3\cdots a_A} \tag{7}$$

which lie in the wanted band, and Q for the remainder

$$Q = \frac{a_{A+1}\cdots a_{r-1}}{a_{A+2}\cdots a_r}\sqrt{a_c} \tag{8}$$

which lie in the transition band. Then, obviously, (4) becomes

$$\tanh\frac{\theta}{2} = PQ. \tag{9}$$

Requirements on Transition Factors

With these formulæ before us, we are now prepared to attack the problem of meeting the double requirement of linear phase shift in the practical transmission band and infinite loss in the practical attenuation band. Expressed analytically, these requirements are simply

$$\tanh\frac{\theta}{2} = \tanh i\frac{\pi f}{2\alpha} = i\tan\frac{\pi f}{2\alpha}, \qquad f < f_A, \tag{10}$$

$$\tanh\frac{\theta}{2} = 1, \qquad f > f_B, \tag{11}$$

[13] It should be noted that, except for the cut-off factor a_c, (4) and (5) are entirely independent. That is, the frequencies $f_1, \ldots f_r$ may be chosen as we desire, in order to control the transfer constant, without in any way affecting the image impedance; and $f_1', \ldots, f_s'$ can be chosen at will without affecting θ. Similarly the constants $\sqrt{K_x/K_y}$ and $\sqrt{K_xK_y}$ may be chosen at will.

where α is a constant which determines the slope of the phase curve.
But it is well known that

$$i \tan\frac{\pi f}{2\alpha} = \frac{i\pi f}{2\alpha}\frac{\left(1-\frac{f^2}{2^2\alpha^2}\right)\left(1-\frac{f^2}{4^2\alpha^2}\right)\cdots}{\left(1-\frac{f^2}{\alpha^2}\right)\left(1-\frac{f^2}{3^2\alpha^2}\right)\cdots}. \tag{12}$$

If, then, in (4) we choose

$$\sqrt{K_x/K_y} = \pi/2\alpha,$$
$$f_1 = \alpha, \qquad f_2 = 2\alpha, \qquad \cdots, \qquad f_A = A\alpha,$$

so that P becomes identical with the first A terms of (12), and if in addition we choose our unit of frequency so that[14] $(A + 1)\alpha = 1$, we readily see that in the transmitted range Q must equal

$$Q = \frac{(1-f^2)\left(1-\frac{f^2}{(1+2\alpha)^2}\right)\cdots}{\left(1-\frac{f^2}{(1+\alpha)^2}\right)\left(1-\frac{f^2}{(1+3\alpha)^2}\right)\cdots}, \qquad f < f_A, \tag{13}$$

while by (9) and (11) in the attenuated range it must be given by

$$\frac{1}{Q} = P = \frac{i\pi f}{2\alpha}\frac{\left(1-\frac{f^2}{2^2\alpha^2}\right)\cdots\left(1-\frac{f^2}{(A-1)^2\alpha^2}\right)}{\left(1-\frac{f^2}{\alpha^2}\right)\cdots\left(1-\frac{f^2}{A^2\alpha^2}\right)}, \qquad f > f_B. \tag{14}$$

Expressed in terms of Gamma functions, (13) and (14) become[15]

$$Q = \frac{\Gamma^4\left(\frac{1}{2\alpha}\right)\Gamma\left(\frac{1-f}{\alpha}\right)\Gamma\left(\frac{1+f}{\alpha}\right)}{\Gamma^2\left(\frac{1}{\alpha}\right)\Gamma^2\left(\frac{1-f}{2\alpha}\right)\Gamma^2\left(\frac{1+f}{2\alpha}\right)}, \qquad f < f_A; \tag{15}$$

and

$$Q = \frac{1}{4\pi i}\frac{1-f}{2\alpha}\frac{\Gamma^4\left(\frac{1}{2\alpha}\right)\Gamma^2\left(\frac{f-1}{2\alpha}\right)\Gamma\left(\frac{f+1}{\alpha}\right)}{\Gamma^2\left(\frac{1}{\alpha}\right)\Gamma\left(\frac{f-1}{\alpha}\right)\Gamma^2\left(\frac{f+1}{2\alpha}\right)}, \qquad f > f_B. \tag{16}$$

[14] This means that we express all frequencies in terms of the first critical frequency of (12) which falls in the transition interval.

[15] The necessary transformations may be found in Whittaker & Watson, "Modern Analysis," 12.13, 12.15, and 12.33.

Asymptotic Series for P and Q

If we now take the logarithm of (15), and apply Stirling's formula, we obtain the asymptotic series

$$\log Q = \log \frac{4^{1/\alpha}\Gamma^4(1/2\alpha)}{8\pi\alpha\Gamma^2(1/\alpha)} + \frac{1}{2}\log (1 - f^2) + \sum_{r=1}^{\infty} \frac{(-1)^r(2^{2r}-1)B_r\alpha^{2r-1}}{2r(2r-1)}\left[\frac{1}{(1+f)^{2r-1}} + \frac{1}{(1-f)^{2r-1}}\right], \quad (17)$$

where the B's are Bernoulli numbers. Since Stirling's formula holds only for $z > 0$, this expansion is valid, as inspection of equation (15) will show, only for $f < 1$, but as the unit of frequency was so chosen that $f_A < 1 < f_B$, this includes the entire wanted band, and none of the attenuating range.

If we apply a similar process to (16) we are again led to (17), except that now the range of validity is $f > 1$. But this includes the entire attenuating range, and none of the wanted band.

That is, the single formula (17) represents the lacunar [16] function Q in both ranges in which it is well defined.

We shall now determine the transition factors $a_{A+1}, a_{A+2}, \cdots, a_c$ by comparison of (8) with (17). If we adopt the notation

$$f_{A+1} = 1 + c_1, \quad f_{A+2} = 1 + c_2, \quad \cdots, \quad f_c = 1 + c_m, \quad (18)$$

so that the c's measure, not the critical frequencies themselves, but their displacement from unit frequency, each factor of (8) has the characteristic form

$$1 - \frac{f^2}{(1+c_j)^2} \equiv \frac{1-f^2}{(1+c_j)^2}\left(1 + \frac{c_j}{1-f}\right)\left(1 + \frac{c_j}{1+f}\right);$$

whence (8) becomes

$$Q = K\frac{\left(1+\frac{c_1}{1-f}\right)\left(1+\frac{c_1}{1+f}\right)\cdots\sqrt{1+\frac{c_m}{1-f}}\sqrt{1+\frac{c_m}{1+f}}}{\left(1+\frac{c_2}{1-f}\right)\left(1+\frac{c_2}{1+f}\right)\cdots\left(1+\frac{c_{m-1}}{1-f}\right)\left(1+\frac{c_{m-1}}{1+f}\right)}\sqrt{1-f^2}, \quad (19)$$

where K is a constant multiplier which depends on the c's. We will neglect it in this analysis since it may be readily determined later from the condition that $Q = 1$ when $f = 0$.

[16] A lacunar function is one which is well defined in several regions, but not capable of analytic continuation from one to the other.

The logarithm of Q is, of course, the sum of the logarithms of the individual factors of this expression. Expanding these as series of powers of $1/(1-f)$ or $1/(1+f)$, and collecting terms of like degree in $1/(1-f)$ and $1/(1+f)$, we obtain

$$\begin{aligned}\log Q &= \frac{1}{2}\log(1-f^2)\\ &\quad+\left(c_1-c_2+c_3-\cdots\pm\frac{1}{2}c_m\right)\left[\frac{1}{1-f}+\frac{1}{1+f}\right]\\ &\quad-\frac{1}{2}\left(c_1^2-c_2^2+c_3^2-\cdots\pm\frac{1}{2}c_m^2\right)\left[\frac{1}{(1-f)^2}+\frac{1}{(1+f)^2}\right]\\ &\quad+\frac{1}{3}\left(c_1^3-c_2^3+c_3^3-\cdots\pm\frac{1}{2}c_m^3\right)\left[\frac{1}{(1-f)^3}+\frac{1}{(1+f)^3}\right]\\ &\quad-\cdots,\end{aligned}\tag{20}$$

where the sign of c_m is plus or minus according as m is odd or even.[17] As the terms of (20) are similar in form to those of (17),[18] we can make the first m terms identical. This leads to the equations

$$\left.\begin{aligned}c_1-c_2+c_3-\cdots\pm\frac{1}{2}c_m &= -\frac{3}{2}B_1\alpha,\\ c_1^2-c_2^2+c_3^2-\cdots\pm\frac{1}{2}c_m^2 &= 0,\\ c_1^3-c_2^3+c_3^3-\cdots\pm\frac{1}{2}c_m^3 &= +\frac{15}{4}B_2\alpha^3,\\ c_1^4-c_2^4+c_3^4-\cdots\pm\frac{1}{2}c_m^4 &= 0,\\ c_1^5-c_2^5+c_3^5-\cdots\pm\frac{1}{2}c_m^5 &= -\frac{21}{2}B_3\alpha^5,\\ \cdots\quad\cdots\quad\cdots &\quad\cdots,\end{aligned}\right\}\tag{21}$$

whose simultaneous solution gives the desired transition factors.

The number m of transition factors used will depend upon the desired approximation to ideal characteristics in the practical transmitting and

[17] When the c's are evaluated it will appear that these series are all absolutely convergent—so that their termwise sum correctly represents log Q—at all positive frequencies outside the interval $(1-c_m, 1+c_m)$. As the complexity of the network is increased, in the approach toward ideal characteristics, the interval of non-convergence closes on the reference frequency 1, and is contained by the given transition interval.

[18] The first term of (17) does not contain f, and may therefore be neglected for the same reason which led us to neglect K in (19).

attenuating ranges and the allowable width of the transition interval. It can best be determined by inspection of results given later.

The result of solving the equations (21) for the ratios c_j/α for values of m between 1 and 5 is given in the following table.

TABLE I
SPACING OF TRANSITION FACTORS

Number of Factors	c_1/α	c_2/α	c_3/α	c_4/α	c_5/α
1	−0.50000				
2	−0.14645	+0.20711			
3	−0.05032	+0.67731	+0.95526		
4	−0.01897	+0.86157	+1.49180	+1.72252	
5	−0.00760	+0.93809	+1.74806	+2.30277	+2.50080

The first of these solutions corresponds to a single frequency, the cut-off, in the transition interval. It follows the uniformly spaced critical frequencies of the practical transmission band at one-half the uniform spacing, α. The other solutions represent networks having, in addition to the cut-off, rational factors which vanish in the transition interval.

When these values of the c's are used in equation (8), with due regard for (6) and (18), the form of Q is completely determined. For example, the frequency pattern corresponding to the case $m = 3$, is illustrated by Fig. 2.

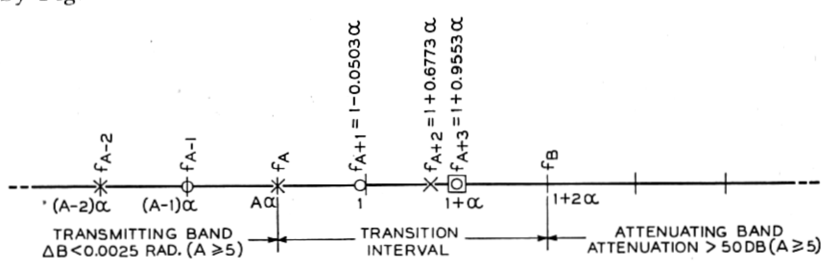


Fig. 2—Location of transition factors with $m = 3$.

Nature of the Approximation

How closely we approach ideal characteristics by this method depends on how nearly log Q is represented by the first m terms of (17) and (20). In both cases the series of omitted terms can be written in the form

$$A_{m+1}\alpha^{m+1}\left[\frac{1}{(1-f)^{m+1}} + \frac{1}{(1+f)^{m+1}}\right] + A_{m+2}\alpha^{m+2}\left[\frac{1}{(1-f)^{m+2}} + \frac{1}{(1+f)^{m+2}}\right] + \cdots, \quad (22)$$

where the A's are constants. In (20) this series is convergent. In (17) it is merely asymptotic. It is known, however, that the error due to ending (17) at any term is numerically less than the first omitted term. Since we are at present interested in small values of α, therefore, we can estimate the error in the approximation from the first term alone.

Inspection of this term shows that the error is greatest in the vicinity of the transition interval, where the factor $1/(1-f)$ is large. It depends upon all three of the quantities f, α and m; but by choosing them in the proper order there is no difficulty in showing that an indefinitely close approximation can be obtained.

The transition interval must first be selected on the absolute frequency scale. It may be as small as we choose. Next, a value of m must be chosen. What value is used is immaterial for our present purposes, although it is important for later applications. Finally, α must be taken small enough so that all transition factors lie in the prescribed transition interval. Otherwise it may be varied at will. But by choosing it small enough, the error of approximation (22) can obviously be reduced without limit for any value of f outside the interval (f_A, f_B). We may thus conclude that only considerations of expense and of manufacturing precision restrict the accuracy of approach to the ideal filter.

For purposes of future reference approximate formulæ for the attenuation and phase in the limiting condition are given below:

$$e^{-A} \doteq -\frac{1}{2} A_{m+1}\alpha^{m+1}\left[\frac{1}{(1-f)^{m+1}} + \frac{1}{(1+f)^{m+1}}\right], \tag{23}$$

$$B \doteq \frac{\pi f}{\alpha} + A_{m+1}\alpha^{m+1}\left[\frac{1}{(1-f)^{m+1}} + \frac{1}{(1+f)^{m+1}}\right]\sin\frac{\pi f}{\alpha}. \tag{24}$$

It will be seen that the attenuation rises monotonically as we recede from the transmission band while the phase curve ripples about the ideal straight line in a sinusoid of varying envelope. The ripples, of course, increase in frequency as α is diminished but since the exponent $m+1$ is always at least 2 they flatten out so rapidly that $dB/d\omega$ approaches constancy nevertheless. We may also observe that, although the absolute time of delay increases indefinitely as α decreases, it varies only as $1/\alpha$, whereas the precision of approximation can be made indefinitely great by choosing m large.

Filters of Other Types

While the preceding analysis has been restricted formally to low-pass filters, its application to filters of other transmission types is a simple matter. We need merely repeat, in each transmission or transition interval, the rules for frequency spacing we have already developed.

The method can be understood from the study of a linear phase shift high-pass filter. Since linear phase shift demands arithmetic spacing of critical frequencies in the transmission band, it is clear that the desired characteristic cannot be obtained over the complete transmission band of the high-pass filter with a *finite* network. This difficulty will, however, be ignored for the moment. A method of modifying the analysis to give a finite filter having a linear phase characteristic in a finite interval above the cut-off will be described later.

We begin, then, by assigning to the transmission band of the structure an infinite, evenly spaced chain of critical frequencies, as in (13). The group of transition factors must evidently simulate the reciprocal of this value in the attenuation range, if the condition of high loss is to be realized; while in the transmission band, they must simulate the P of our earlier analysis if we are to obtain a linear phase characteristic. These conditions would be met by using for our transition function the reciprocal of (19), using for the c's the same values as before. Such a group, however, is not physically realizable as part of a high-pass transfer constant, since the rational factors would occur outside the theoretical transmission band. If, however, we transfer the factor $(1 - f^2)$ from (13) to (14), and seek a new Q whose values will take the reciprocals of the old, thus altered, we obtain a series identical with (17) except for a change of sign in every term but $\frac{1}{2} \log (1 - f^2)$. This change, however, reverses the sign of the right-hand members of (21), and therefore changes the sign of each c. The new solution then is the same as the original solution except that the factors occur in reverse order on the frequency scale. They can thus appropriately be combined with the remaining portion of the high-pass transfer constant expression.

A linear phase shift band-pass filter can be constructed similarly. The groups of transition factors associated with the upper and lower cut-offs should follow the arrangements prescribed, respectively, for low-pass and high-pass filters. An illustration will be found in Part II.

The Impedance Property

It will be recalled that the problem of approximating the ideal transmission characteristics for each type of filter was solved only on the assumption that the image impedance could be adjusted to a nearly uniform value in the practical transmitting band. We can now quickly show how the desired impedance is to be obtained. It is merely necessary to observe that for any filter there exists a complementary structure with the same arrangement of critical frequencies, but having the transmitting and attenuating bands interchanged. The complementary structure is found by replacing the Z_y branch of the original lattice by the inverse impedance $Z_y' = R^2/Z_y$. When these are substituted in (1) and (2), the new transfer constant, θ', is found to be

$$\tanh\frac{\theta'}{2} = \sqrt{\frac{Z_x}{Z_y'}} = \frac{1}{R}\sqrt{Z_x Z_y} = \frac{1}{R} Z_I,$$

and the new image impedance, Z_I',

$$Z_I' = \sqrt{Z_x Z_y'} = R\sqrt{\frac{Z_x}{Z_y}} = R \tanh\frac{\theta}{2}.$$

Thus, for any filter, the problem of adjusting the image impedance to the constant R in its transmitting band is the same as the problem of adjusting $\tanh \theta/2$ to 1 in the attenuating band of the complementary filter. The latter problem, however, is merely a restatement of our original requirement of high loss in attenuating bands and has already been studied for various types of filters.

It follows from this relation that the transfer constant expressions which are appropriate for low-pass and band-pass filters furnish suitable solutions for the impedance problem in high-pass and band-elimination structures. We might also use our high-pass transfer constant expression as a low-pass impedance characteristic except for the difficulty previously mentioned that it requires an infinite number of elements. This difficulty can be avoided, however, by observing that by interchanging coils and condensers we can convert any low-pass filter into a high-pass structure having the same characteristics on a reciprocal frequency scale. We can thus use the finite low-pass solutions to obtain the required finite high-pass filter having high attenuation. For example, if we begin with a low-pass filter having three evenly spaced critical frequencies and a half spaced cut-off the resulting critical frequencies (including the cut-off) are in the ratio $1 : 7/6 : 7/4 : 7/2$.

The device of inverting the frequency scale is, of course, not available to produce a finite high-pass filter having linear phase shift throughout its transmission band since the linear phase property is thereby destroyed. It can be used, however, to produce a finite filter having linear phase shift for a limited region above its cut-off.

To see this, it is merely necessary to observe that the set of rational factors appearing in the low-pass image impedance expression described in the preceding paragraph must approximate the reciprocal of the cut-off factor at lower frequencies. We can therefore use such a set of factors to replace the upper cut-off factor of a band-pass filter, obtaining thereby a high-pass structure which approximates the ideal characteristics over a portion of the transmitting band.

If the cut-off factor of the low-pass filter transfer constant be similarly replaced by rational factors, there results an all-pass "delay network" having a constant impedance and a phase characteristic linear below the original cut-off frequency. This network is of particular interest for its relation to the classic problem of the simulation of a smooth line. As it stands, the network evidently simulates an ideal dissipationless line. To include the effects of dissipation we need merely add resistance and leakance to the coils and condensers in the proportions in which they occur in the actual line.

Part II—Design of Practical Filters

Thus far we have been interested primarily in demonstrating that an indefinitely close approximation to the ideal characteristics could be obtained when all restrictions with respect to economy of elements were removed. In practical designs, on the other hand, we wish to approximate the ideal characteristics only within moderate limits, and our interest centers upon the choice of the most economical network which will prove satisfactory. We must now reappraise the theory from this point of view.

One question which must be examined is that of determining values for m and α which will result in the most economical network meeting a prescribed standard of performance. A second is concerned with the possibility of changing the nature of the approximation with respect either to the frequency, or to the relative emphasis laid upon the phase and attenuation characteristics. In many practical designs such changes can be obtained by slight modifications of the theoretical design parameters and lead to corresponding economies in the use of elements. In investigating both questions we must remember that since α is no longer necessarily small, as it was in the theoretical analysis, the frequency interval actually occupied by the transition factors may be appreciable. Consequently it becomes important to investigate the behavior of the network in this part of the frequency range with more care than was hitherto necessary.

The variety of possible design requirements precludes the possibility of a thorough analytic treatment of these questions. The choice of the most economical network meeting given requirements consequently cannot always be made without trial. The procedure may, however, be considerably facilitated by a study of the curves and illustrative material given in the sections which follow. The first two sections show the quantitative relations to be expected when the theoretical design parameters are adhered to strictly. The remaining sections indicate modifications obtainable by making slight changes in the theoretical parameters.

Approximate Computation of Network Characteristics

When the frequency in which we are interested is not too close to the transition interval, an approximate determination of the phase and attenuation characteristic is most easily made from (23) and (24). The A's appearing in these expressions are shown in the accompanying table.[19] In addition to A_{m+1} the table also supplies values of A_{m+2}

TABLE II

COEFFICIENTS IN SERIES EXPANSIONS FOR APPROXIMATION ERRORS IN PHASE AND ATTENUATION CHARACTERISTICS

m	1	2	3	4	5
A_{m+1}	−0.063	−0.044	−0.051	−0.084	−0.17
A_{m+2}	−0.063	+0.00011	+0.10	+0.41	+1.52
A_{m+3}	−0.0078	+0.050	−0.047	−1.07	−7.60

and A_{m+3}, for use if additional terms in the general expression (22) are desired.

A study of equations (23) and (24) shows that, aside from the constant factor A_{m+1}, each expression can be resolved into two factors by means of which the contributions of the various design parameters can be somewhat segregated. The first factor, α^{m+1}, is chiefly important in determining the effect of various choices of α and m on the approximation error, while the factor $\left[\frac{1}{(1-f)^{m+1}}+\frac{1}{(1+f)^{m+1}}\right]$ expresses the variation of the network characteristics with frequency. In order to facilitate design work the quantity

$$-20\log_{10}\frac{A_{m+1}}{2}\left[\frac{1}{(1-f)^{m+1}}+\frac{1}{(1+f)^{m+1}}\right]$$

[19] In preparing the table, coefficients of corresponding terms in the series expansions for (17) and (20) have been combined, so that the coefficients as given represent the accumulated errors of both approximations.

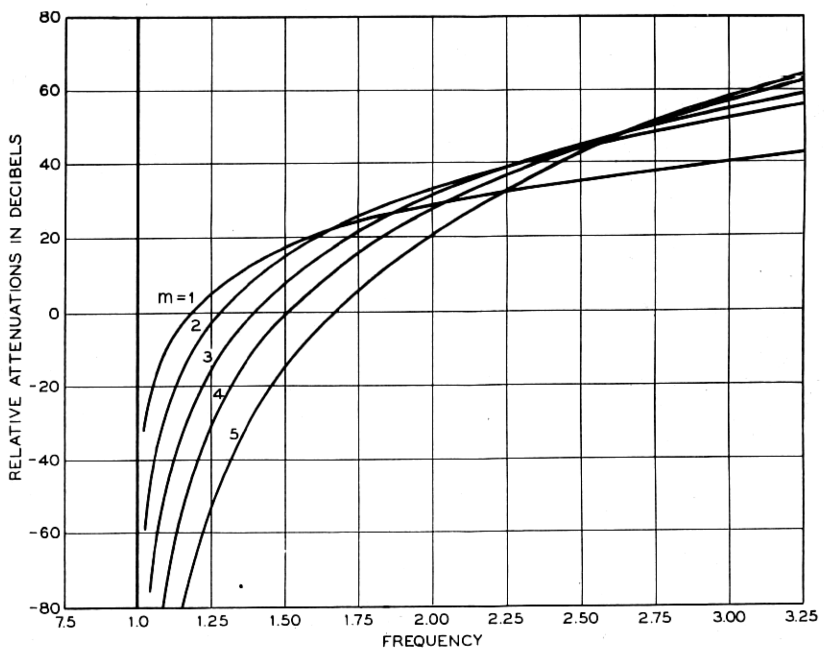


Fig. 3—Chart for loss computations.

has been computed for values of $f > 1$ and is shown plotted for various m's in Fig. 3. The approximate attenuation, in db, for any given values of α and m can be obtained from the chart by adding $20(m + 1) \log_{10} 1/\alpha$ to the appropriate curve.

A similar chart for the phase characteristic is furnished by Fig. 4, which represents the quantity $\frac{180A_{m+1}}{\pi}\left[\frac{1}{(1-f)^{m+1}} + \frac{1}{(1+f)^{m+1}}\right]$. The values given are arithmetic, although the scale is logarithmic. The approximate envelope of the ripple in the phase characteristic about the ideal straight line can therefore be found, in degrees, by multiplying the chart values by α^{m+1}.

In using these charts it should be remembered that they are based upon the approximate formulæ (23) and (24) which fail in the vicinity of the transition interval. The results, therefore, should always be checked by an exact computation.[20] It should also be observed that in complicated filters the numerical departure of tanh $\theta/2$ from its ideal value in most frequency ranges is very small. The effects of slight errors in calculations or of small deliberate variations in the

[20] See, for example, the comparisons in Figs. 11 and 12.

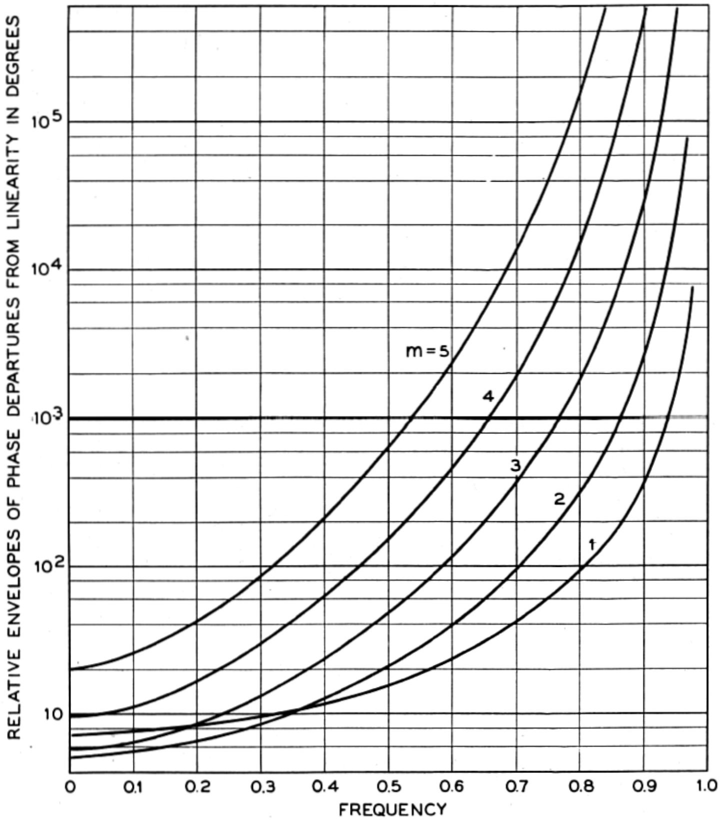


FIG. 4.—Chart for phase computations.

design parameters, may therefore be correspondingly important.[21] Since slight adjustments in the design parameters will normally occur, these charts are chiefly of value in making preliminary estimates.

It is apparent that the approximation error at a given frequency can be diminished either by increasing m or reducing α. Element for element, an increase in m is much the more powerful method. Since the total number of elements in the network is nearly proportional to

[21] A simple example is furnished by the choice of the numerical constant multiplying $\tanh \theta/2$ as a whole. It will be remembered that the constant was left undetermined in the solution for the c's. In the original equation (14) it was chosen to give the best characteristics in the neighborhood of $f = 0$. In preparing Figs. 3 and 4, on the other hand, it was chosen with reference to the characteristics near $f = \infty$, since the error expression used in these figures vanishes at that point. The two conditions are very nearly equivalent; as we can see, for the half-spaced cut-off solution at least, by means of Wallis' theorem. Since they are not identical, however, a change from one to the other may produce a relatively large, though practically unimportant, effect at extreme frequencies.

$m + (1/\alpha)$ it would therefore appear that the most economical structure meeting given requirements will be obtained by using a large m in combination with a large α. This procedure is, however, restricted by two considerations. The first is chiefly theoretical. Since the series we have been using is merely asymptotic, the successive terms obtained by choosing progressively higher m's eventually grow larger. For ordinary values of α, however, the value of m at which the series begins to diverge lies beyond the range of practical interest. A more important limitation is the fact that as we increase the number of transition factors, the width of the transition interval, as measured in terms of α, also increases. Thus, the spread between the last uniformly spaced critical frequency and the cut-off, which is $\alpha/2$ for $m = 1$ and about 2α for $m = 3$, has risen to more than 3.5α for $m = 5$. In each case a certain additional allowance is of course required for the region of rising attenuation beyond the cut-off. When the transition interval is fixed on an absolute frequency scale, therefore, the permissible values of m will depend upon the choice of α. Unless the transition interval is unusually broad only low values of m will be of practical interest.

Illustrative Characteristics

The curves shown in Figs. 3 and 4 are not of use in the neighborhood of the transition interval. To supplement them, therefore, exact computations on a number of typical structures have been made. One set was obtained by choosing $\alpha = 1/12$ and computing the characteristics corresponding to various m's. The resulting phase characteristics are shown by Fig. 5. Since the departures from linearity are too small to be noticeable when the characteristics as a whole are drawn, the figure shows only the departures themselves in terms of an envelope similar to that used for Fig. 4. The curves are drawn approximately as far as the last evenly spaced critical frequency which marks the practical limit of the range within which a high degree of phase linearity is to be expected. Since the curves vary rapidly in this vicinity, however, the fact that they are merely envelopes is important in determining the exact performance of the structure. Curves of the phase characteristics in the transition interval will be given later.

The attenuation characteristics are shown by Fig. 6. As m is increased, the cut-off moves to successively higher frequencies because of the progressively broader intervals consumed by the transition factors. Once past the cut-off, however, the curves for large values of m rise more rapidly and quickly cross the others.

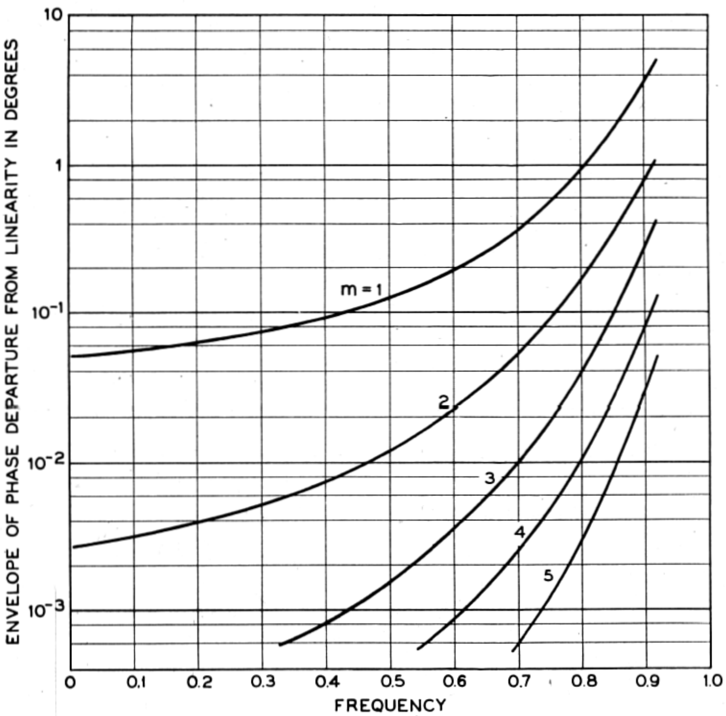


Fig. 5—Low pass filters with $\alpha = 1/12$. Envelopes of phase departures.

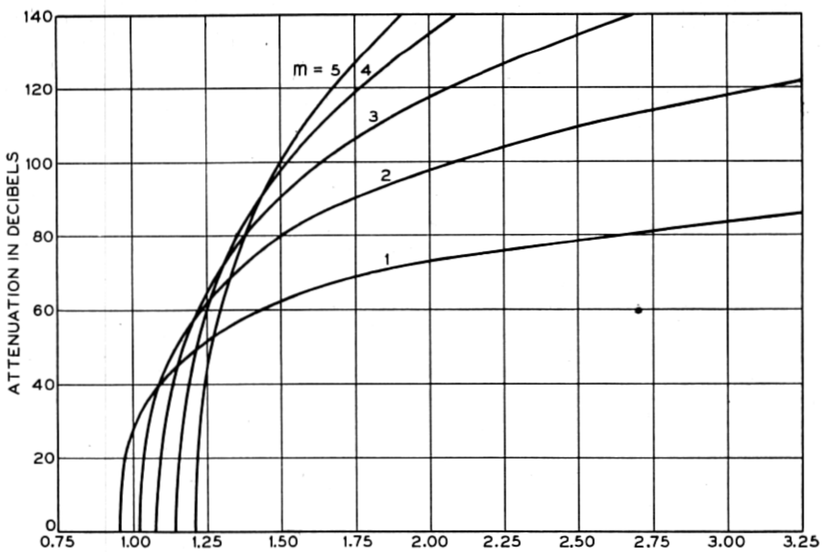


Fig. 6—Low pass filters with $\alpha = 1/12$. Attenuation.

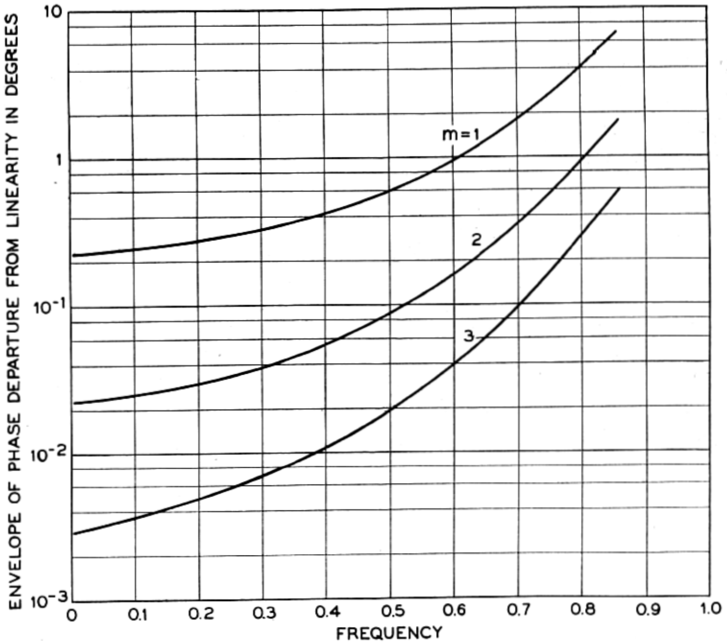


Fig. 7—Low pass filters with $\alpha = 1/6$. Envelopes of phase departures.

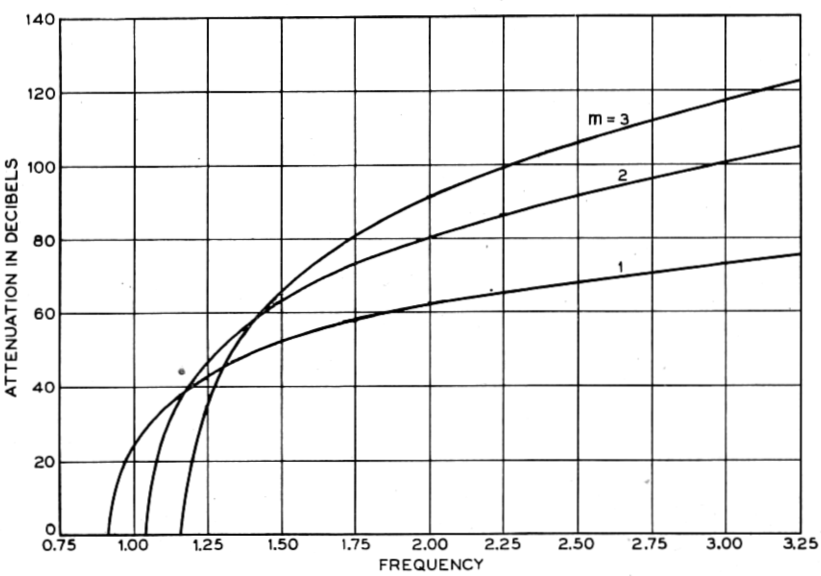


Fig. 8—Low pass filters with $\alpha = 1/6$. Attenuation.

A second set of characteristics was obtained by choosing $\alpha = 1/6$ and adding various groups of transition factors in a similar fashion. The results are shown by Figs. 7 and 8. The characteristics are drawn only for m's between 1 and 3 in this case, since with larger m's the

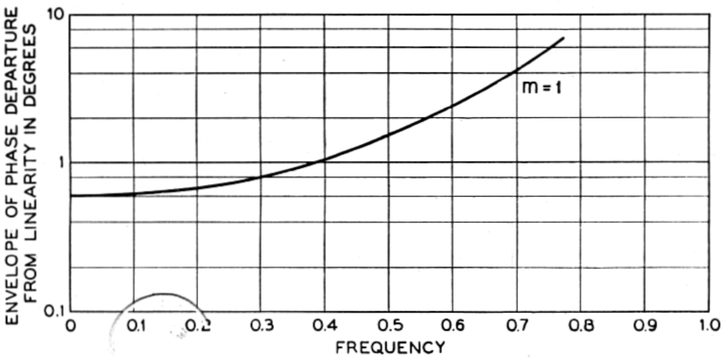


Fig. 9—Low pass filter with $\alpha = 1/4$. Envelope of phase departures.

transition interval becomes disproportionately wide in comparison with the practical transmission range. Still a third set, corresponding to $\alpha = 1/4$ and $m = 1$ is shown by Figs. 9 and 10.

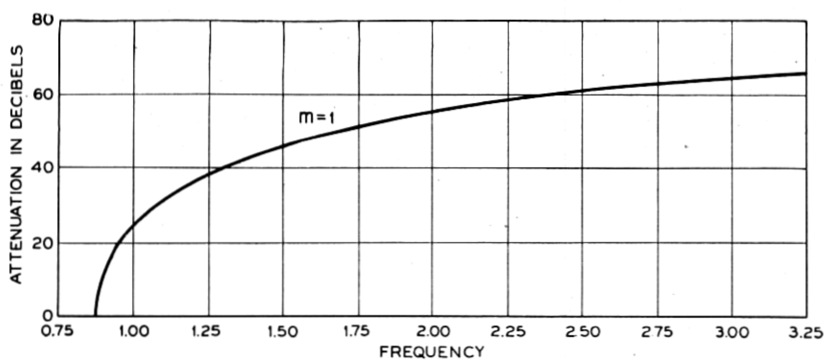


Fig. 10—Low pass filter with $\alpha = 1/4$. Attenuation.

As an illustration of the accuracy to be expected from the approximate method, a comparison between the results obtained by this method and the exact characteristics is shown in Figs. 11 and 12 for the cases $m = 1$ and $m = 2$ of Figs. 5 and 6. On the logarithmic scales used for the figures, the curves appear to be in good agreement almost up to the transition interval. The actual numerical departures in the vicinity of that interval, however, are quite large.

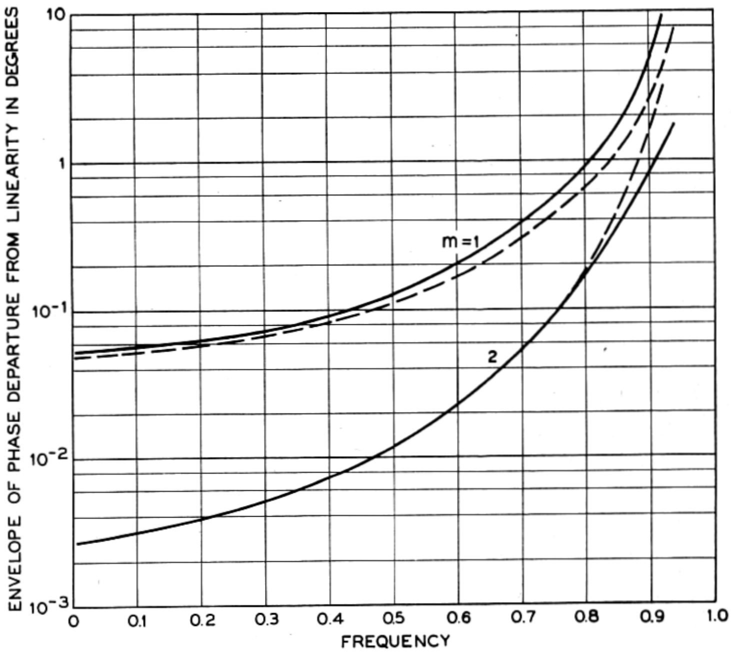


Fig. 11—Approximate and exact envelopes of phase departure for low pass filters with $\alpha = 1/12$.

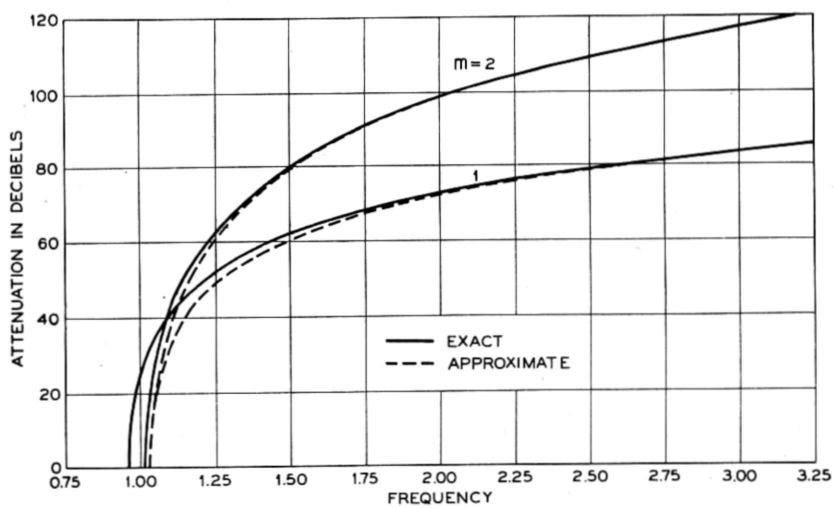


Fig. 12—Approximate and exact attenuation characteristics for low pass filters with $\alpha = 1/12$.

Image Impedance Characteristics

In virtue of the relationship previously developed between the image impedance of a given filter and the transfer constant of its complement, the curves just given might also be used to determine the impedance characteristic. However, the precision required in the approximation of Z_I to R in practical filter design is much less than that required in the approximation of tanh $\theta/2$ to unity. A satisfactory characteristic can therefore be obtained with a much smaller number of critical frequencies. In a low-pass filter, for example, one or two impedance controlling frequencies is usually sufficient. With such a small number of critical frequencies the analytical machinery we have set up is unnecessarily cumbersome. The problem can be solved more effectively by simple cut and try methods, or by the methods advanced by Cauer [7] and Zobel.[22] For the sake of completeness, however, several illustrative characteristics are given in Fig. 13. They correspond to the choice of impedance controlling frequencies

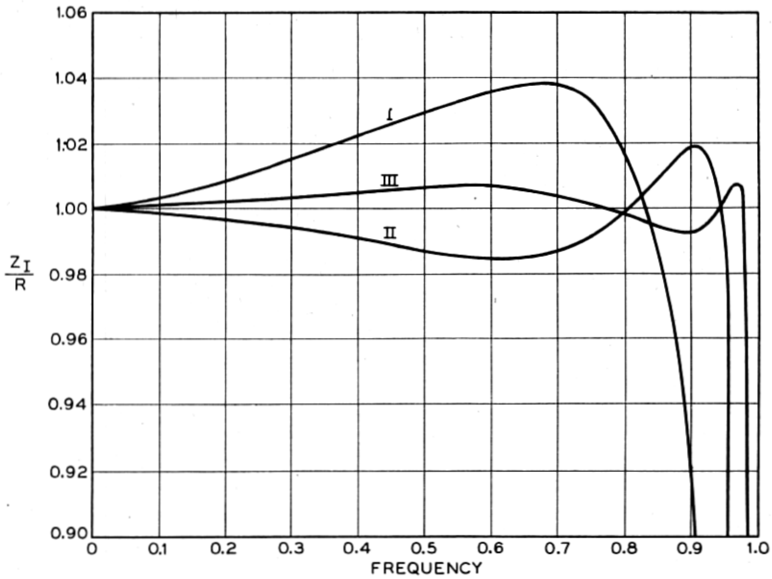


Fig. 13—Typical low pass filter image impedance characteristics.

[7] "Siebschaltungen," loc. cit.

[22] This *Journal*, Apr. 1931, p. 284. Zobel's work is not stated in terms of the lattice parameters. A simple m-type termination (of low-pass or high-pass type) can be identified with a lattice image impedance having one impedance controlling frequency while an mm'-type termination can be identified with a lattice impedance having two such frequencies. The numerical values he gives can therefore readily be adapted to the lattice design problem.

shown in Table III. An illustration of the results obtainable with the present method using a large number of critical frequencies, is furnished by Fig. 14, since the curve can evidently be interpreted as a representation of Z_I/R for a certain high-pass filter.

TABLE III

IMPEDANCE CONTROLLING FREQUENCIES CORRESPONDING TO CHARACTERISTICS OF FIG. 13

I	II	III
$1.250 f_c$	$1.048 f_c$ $1.448 f_c$	$1.013 f_c$ $1.096 f_c$ $1.584 f_c$

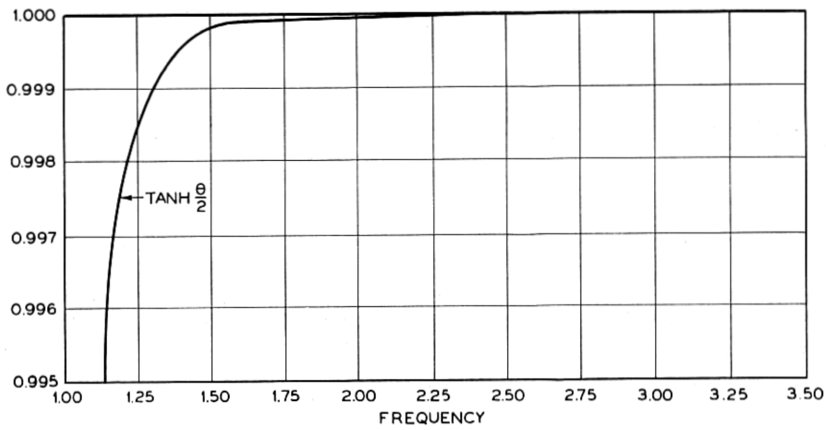


Fig. 14—Tanh $\theta/2$ for low pass filter with $m = 2$, $\alpha = 1/12$.

Weighting the Approximation

In the limiting case in which α is very small while m is fixed, the methods we have followed give the best obtainable results with respect to both attenuation and phase, for the errors in both characteristics depend upon higher order terms and become negligible as α approaches zero. In a practical design, for which α is finite, on the other hand, it will frequently be desirable to make slight adjustments in such parameters as the transition factors or the constant multiplier of the tanh $\theta/2$ expression in order to take some account of the higher order terms. The effect may be either to improve both the phase and attenuation characteristics or, more usually, to improve one at the expense of the other.

The general nature of the problem is illustrated by Fig. 14, which represents a sketch of tanh $\theta/2$ corresponding to the $m = 2$ curve of Fig. 6. It will be seen that the curve rises monotonically toward the line unity, at which $\theta = \infty$. What we should evidently like to obtain by slight alterations in the design parameters is a curve which rises more rapidly, or perhaps one which ripples about unity. It is also evident that the curve approximates unity so closely that even slight adjustments may produce a radical effect. To take the simplest possibility, if the constant multiplier of tanh $\theta/2$ is slightly increased, so that the curve crosses unity at a finite frequency, the appearance of the resulting attenuation characteristic will be greatly altered. The net gain in the general level of attenuation secured, however, will be not more than 6 db. Similar remarks might be made with respect to the phase characteristic.

The relation between the phase and attenuation characteristics where such adjustments are made can be illustrated most easily by reference to the elementary half-spaced cut-off solution for the transition factors. It will be recalled that this solution was obtained by equating the coefficients of the first powers of $1/(1 - f)$ and $1/(1 + f)$ in (17) and (20). The approximation error thus depends chiefly upon the succeeding term involving $1/(1 - f)$ and $1/(1 + f)$ to the second power. A study of the expression shows that the error makes Q too small in both the transmitting and attenuating ranges. If the phase characteristic is the more important this error can be partly compensated by slightly increasing the normal half-space between the cut-off and the preceding critical frequency. On the other hand, the attenuation will be improved if the interval between the cut-off and the preceding critical frequency is decreased. To a more limited extent, both characteristics can be improved by increasing the constant factor which multiplies tanh $\theta/2$ as a whole.

A similar study might be made of the other groups of transition factors, although the discussion would naturally become more complicated. In general it appears, as with the half-spaced cut-off solution, that the attenuation characteristic will be improved by a slight decrease in the spacings of the transition factors, while the phase characteristic will be improved if they are slightly increased. It should be remarked, however, that as the network becomes more complicated, either by a reduction in α or an increase in m, the desirable modifications in the theoretical spacings are reduced. This becomes evident if it is recalled that the transition factor spacings are proportional to α while the error is roughly proportional to α^{m+1}. It is therefore to be expected that the appropriate modifications in the spacings between transition frequencies will be of the order of magnitude of α^m times their original values.

The relationship between the phase and attenuation characteristics can be seen in another light if we observe that the improvement in attenuation which comes from the use of several transition factors is due essentially to a progressive decrease in the interval between critical frequencies as the cut-off is approached. In the final solution, for example, the intervals between critical frequencies are initially almost equal to the constant interval α. Thus in this solution, the interval between f_A and f_{A+1} is 0.992α and that between f_{A+1} and f_{A+2} is 0.945α. As the cut-off is approached, however, the interval gradually decreases to about 0.2α. In the transition interval, consequently, the phase characteristic is originally almost linear and curves upward sharply near the cut-off. Thus if the phase requirement is not severe we can consider that the first part of the transition region falls within the practical transmitting band, thereby securing a better attenuation characteristic than would be possible if the spacing of critical frequencies in the transmitting band were strictly uniform. A sketch of the phase characteristics through the transition interval

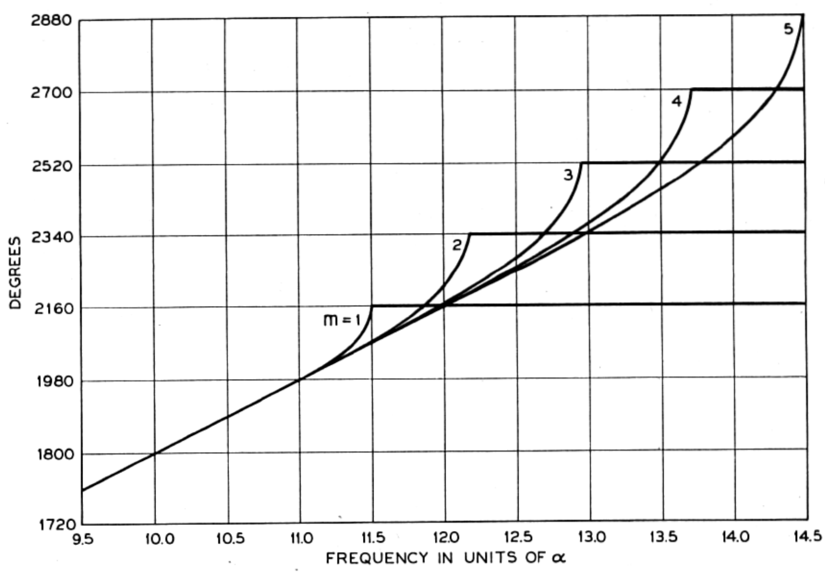


Fig. 15—Transfer constant phase shift in the transition interval; $\alpha = 1/12$.

for the networks corresponding to Figs. 5 and 6 is shown by Fig. 15. The last evenly spaced critical frequency falls at 11α.

In the extreme case when no phase requirement is imposed, it is reasonable to expect that the best attenuation characteristic will be obtained if the progressive reduction in the spacing between critical frequencies extends over the complete transmission band, so that the phase characteristic should resemble that of familiar ladder type filter structures by becoming continually steeper as the cut-off is approached. The exact arrangement will, however, depend upon the desired type of best approximation to perfect suppression. If the approximation is to be best at frequencies most remote from the cut-off, the critical frequencies must be evenly spaced along an ordinary arc sine curve. In the Tschebycheffian type of approximations studied by Cauer, on the other hand, the spacing must be uniform along the arc of a certain sn function.

Design of a Band-Pass Filter

To illustrate the manner in which small modifications of the theoretical frequency spacings may be employed to control the relative emphasis placed on the phase and attenuation characteristics, we may consider the design of a practical band-pass filter. Suppose that the practical transmitting band is the 2,250-cycle interval between 11,375 and 13,625 c.p.s., in which the approximation of the phase characteristic to linearity is specified by the requirement that $\partial B/\partial \omega$, the so-called "delay," deviate from its average value by less than 0.1 millisecond. The transition intervals are 500 c.p.s. each, beyond which the loss is to be not less than 50 db.

The comparatively liberal tolerances suggest that the approximation furnished by $m = 1$ will be adequate. We notice that we can fit 10 uniform intervals of 250 c.p.s. between 11,250 and 13,750 c.p.s., which locates the half-spaced cut-offs at 11,125 and 13,875 c.p.s. respectively. When the characteristics corresponding to this design are checked, it is found that the phase characteristic is rather better than required, while the loss characteristic is weak.

We then turn to the solution with $m = 2$, making a compensating reduction in the number of uniform intervals. The critical frequency allocation for this case is shown in Table IV. This arrangement meets

TABLE IV

CRITICAL FREQUENCY ALLOCATION FOR LINEAR PHASE SHIFT BAND-PASS FILTER

$m = 1$	$m = 2$	Modified	Limits of Required Linear Region
11,125	11,198	11,174	
11,250	11,289	11,265	11,375
11,500	11,500	11,500	
....	....	....	
13,500	13,500	13,500	13,625
13,750	13,711	13,735	
13,875	13,802	13,826	

the loss requirement with a large margin of safety, but the phase shift curve departs seriously from linearity near the last useful frequencies, which fall in the shortened intervals of the transition factors.

With these two attempts as guides, a compromise frequency pattern which exactly suits the conditions of the problem is readily arrived at. In contrast to the transition factor spacings of 0.853α and 0.353α, as shown by the solution for $m = 2$, those actually adopted are 0.94α and 0.375α so as to make the first transition spacing more nearly uniform with those in the pass-band. The indicated frequency pattern is shown in the third column of the table.

As the values of these transition factors near the band edges are somewhat too large they lead to larger undulations of the phase characteristic in those regions than near the band center. The approximations can be rendered more uniform throughout the band without serious consequence to the loss characteristic by multiplying the tangent expression by a constant slightly smaller than unity. In this case the value chosen was $K_1 = 0.9975$.

The final "delay" and loss characteristics, corrected for the effects of dissipation, are exhibited by Figs. 16 and 17. A noteworthy result

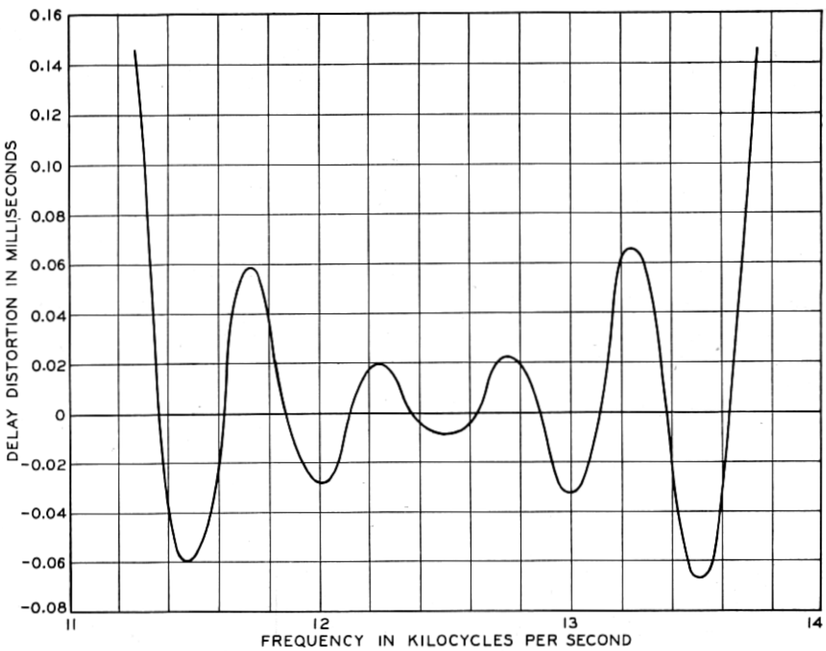


Fig. 16—Variation in the phase slope of a band pass filter.

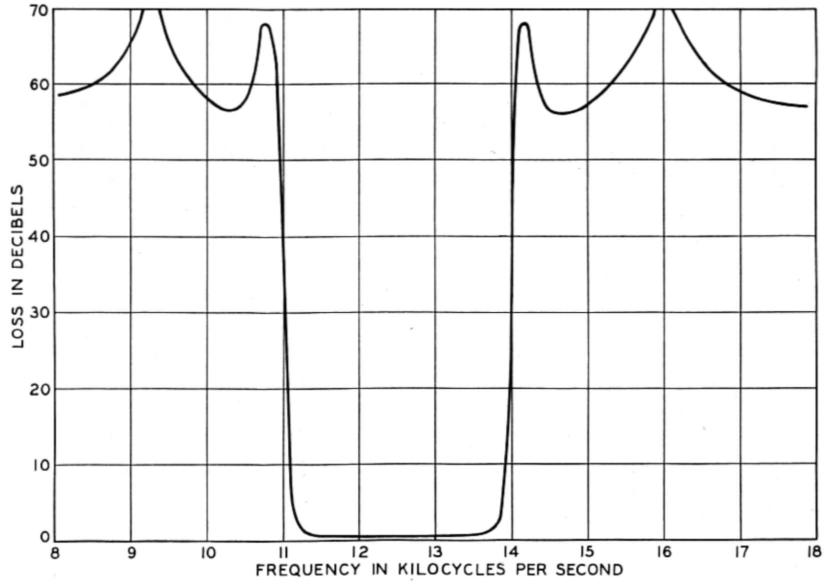


Fig. 17—Attenuation characteristic of a band pass filter.

of modifying the theoretical spacings for the transition factors has been to introduce peaks of loss near the band edges. The shortened intervals adjoining the cut-offs produce tanh $\theta/2$ curves which rise rapidly beyond these points to maxima slightly greater than unity, instead of approaching unity monotonically.

Having thus located the critical frequencies, we may readily complete the design of the filter in lattice form.

The formulation of the transfer-constant expression results in

$$\tanh\frac{\theta}{2} = K_1 \frac{\sqrt{1-\frac{f^2}{f_a^2}}\left(1-\frac{f^2}{f_2^2}\right)\cdots\left(1-\frac{f^2}{f_{10}^2}\right)\sqrt{1-\frac{f^2}{f_b^2}}}{\left(1-\frac{f^2}{f_1^2}\right)\left(1-\frac{f^2}{f_3^2}\right)\cdots\left(1-\frac{f^2}{f_9^2}\right)\left(1-\frac{f^2}{f_{11}^2}\right)},$$

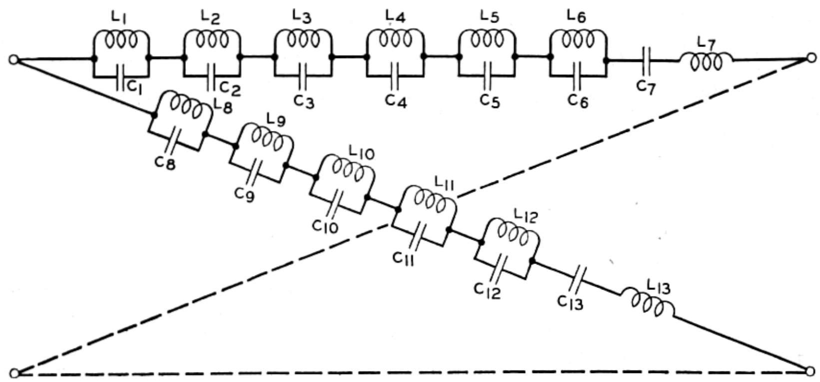
where f_a and f_b represent the cut-offs, and the other f's intervening critical frequencies in order of magnitude, as shown by Table IV.

A suitable form for the image impedance must next be obtained and since it is normally determined by requirements with which we are not now concerned, we will adopt the simplest possible expression, namely

$$Z_I = K_2 \frac{\sqrt{1-f^2/f_a^2}\sqrt{1-f^2/f_b^2}}{if},$$

where K_2 is determined from the condition that $Z_I = 600$ ohms when

$f = \sqrt{f_a f_b}$. The impedance functions Z_x and Z_y are now readily found by means of (1) and (2), and with the help of Foster's formula the element values can be obtained. These are shown in Fig. 18.



$L_1 =$ 0.0675 mh.	$C_1 =$ 3.035 mf.
$L_2 =$ 0.6529 mh.	$C_2 =$ 0.2810 mf.
$L_3 =$ 0.1958 mh.	$C_3 =$ 0.8622 mf.
$L_4 =$ 0.1770 mh.	$C_4 =$ 0.8801 mf.
$L_5 =$ 0.1330 mh.	$C_5 =$ 1.085 mf.
$L_6 =$ 0.0404 mh.	$C_6 =$ 3.325 mf.
$L_7 =$ 35.66 mh.	$C_7 =$ 0.0046 mf.
$L_8 =$ 0.1620 mh.	$C_8 =$ 1.182 mf.
$L_9 =$ 0.2021 mh.	$C_9 =$ 0.8703 mf.
$L_{10} =$ 1.335 mh.	$C_{10} =$ 0.1214 mf.
$L_{11} =$ 0.1097 mh.	$C_{11} =$ 0.9112 mf.
$L_{12} =$ 0.1089 mh.	$C_{12} =$ 1.276 mf.
$L_{13} =$ 35.86 mh.	$C_{13} =$ 0.0046 mf.

Fig. 18—Band pass filter.

This example illustrates the way in which the analysis may be applied to a typical problem in network design. The practical design would not ordinarily be complete at this point, however, since, as was mentioned previously, it is seldom desirable actually to construct the network as a single symmetrical lattice. Improved stability with respect to variations of the elements from their design values is obtained if the lattice is resolved into its components, that is, the elementary lattice sections which when operated in tandem have the same transmission properties. This question is discussed in a recent paper.[3] Furthermore, unbalanced structures equivalent to the symmetrical lattice but employing fewer elements are known,[23] and expense can usually be reduced by resorting to one of these.

[3] H. W. Bode, loc. cit. It may be interesting to observe that in the terminology of that paper the elementary constituents of linear phase shift filters are usually complex *m* sections.

[23] A linear phase shift lattice filter cannot, of course, be constructed as a sequence of Π or T sections, but equivalences in generalized bridged-T configurations exist. General equivalences in configurations employing ideal transformers are familiar in the literature. See, for example, Cauer, loc. cit., or Jaumann, *E. N. T.*, July, 1932, p. 243.

PART III—FILTERS WITH LINEAR PHASE SHIFT THROUGH THE CUT-OFF

It was the conclusion of the theoretical discussion that any desired approximation to ideal filter characteristics may be obtained from a finite network, so long as a finite transition interval separates transmitting from attenuating bands. The transition interval can be taken small at pleasure, but very small transition intervals are associated with networks of many natural frequencies and numerous elements. We have already seen how considerable economies in meeting a given attenuation requirement could be obtained if the phase requirement were subordinated or removed entirely. We now consider the contrary case, in which major emphasis is placed upon the phase characteristic of the filter. Filters of this type are of practical interest in picture transmission systems since instruments used in the reproduction of images seem to be much more sensitive to the effects of phase distortion than the ear. The selectivity required from filters used in such systems is comparatively modest, but phase linearity is required not only in the practical transmission band but also through the transition interval into the region of rising attenuation.

In one important particular the present problem differs from those previously considered. In the present analysis we can no longer regard the adjustment of Z_I and the adjustment of θ as independent problems. On the contrary, in the attenuating region the contribution of θ to the phase shift is constant and we must therefore rely upon reflection effects to maintain the desired linear characteristic. Moreover, near the cut-off θ must be very carefully adjusted with respect to Z_I in order that the contributions to the phase shift from reflection and interaction effects may preserve the linearity through the transition band also. The added restrictions imposed by the extension of the phase requirement require a revision of the frequency spacings already found, and set limits upon the approximation to ideal characteristics obtainable from reactive networks of reasonable complexity.

Use of Reflection Effects to Produce Linear Phase

In the practical transmission band, Z_I can be adjusted to approximate R sufficiently closely to make reflection and interaction effects negligible. Therefore, in this range the total insertion phase is the same as the transfer constant phase, and, as before, is to be obtained from a chain of uniformly spaced critical frequencies in $\tanh \theta/2$. In the practical attenuating band, on the other hand, we find that the imaginary part of θ is either 0 or π, while interaction effects can be ignored if we assume the loss to be reasonably high. The variation of the phase shift with frequency must therefore be attributed to reflection effects, which we can write as

$$e^{\theta_r} = \frac{\left(1 + \frac{Z_I}{R}\right)^2}{4\frac{Z_I}{R}},$$

where θ_r is the sum of the reflection effects at the two ends of the structure.

Since Z_I is reactive in the attenuating band, the angle of the denominator in this equation is $\pm \pi/2$, while that of the numerator is 2 arctan Z_I/iR. Thus

$$B_r = \mp \frac{\pi}{2} + 2 \arctan \frac{Z_I}{iR}. \tag{25}$$

Except for the constant term, which we will consider presently, this is a function of precisely the type we have been considering. Hence if the impedance controlling factors are spaced at the same uniform interval that was used in the pass-band, the phase slope will be constant and equal in both bands.

Phase Characteristics in Transition Intervals

The transition factors—or rather, factor, since clearly we have to rule out the solutions for $m > 1$—must be determined so that these linear parts of the phase characteristic are joined by a chord of the same slope. If we suppose the transition interval to be bounded by the last uniformly spaced frequencies of the transfer constant and image impedance chains, and to contain only the cut-off factor, it is easily shown that it must include a net change in phase of $3\pi/2$ radians. The interval must therefore contain 3/2 uniform spaces if the average slope is to be correct. Considerations of symmetry to be described later require that the cut-off be the center of the interval, which thus comprises two three-quarter spaces. The behavior of the several components of the total insertion phase is exhibited by Fig. 19, in which B, B_r, and B_i refer respectively to the phase shifts contributed by the transfer constant, by reflection effects, and by the interaction factor. The mutually annulling discontinuities of $\pi/2$ radians in B_i and B_r at the cut-off are noteworthy.

The fact that this choice of parameters is sufficient as well as necessary to obtain the desired linearity of phase shift is not easily shown analytically. It can, however, be verified by direct computa-

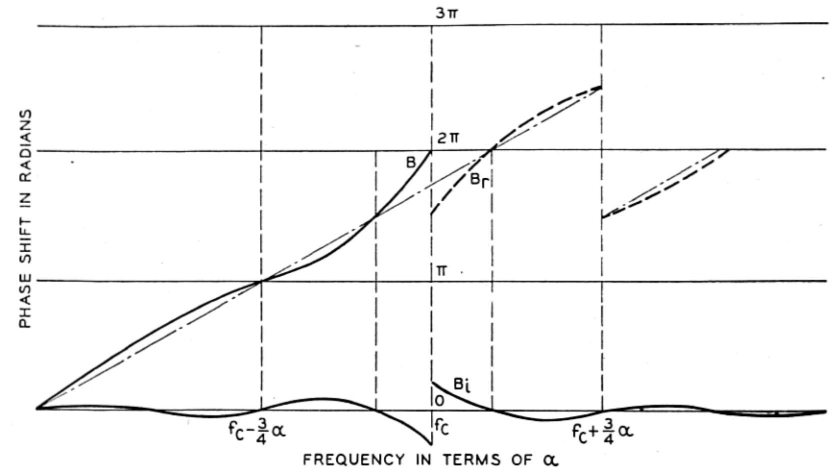


Fig. 19—Transfer, reflection, and interaction phase in the transition interval.

tion. For this purpose the customary resolution of the total insertion loss into transfer constant, reflections, and interaction is not very useful because of the indeterminacies found at the cut-off. This difficulty is avoided by expressing Z_I and θ in terms of the lattice impedances, in which event

$$e^{\gamma} = \frac{1 + \dfrac{Z_x Z_y}{R^2} + \dfrac{Z_x}{R} + \dfrac{Z_y}{R}}{\dfrac{Z_x}{R} - \dfrac{Z_y}{R}}, \tag{26}$$

where γ is the total insertion loss.

If iX_x and iX_y be written for Z_x and Z_y, the insertion loss and phase shift are given by

$$\tan B_{\gamma} = \frac{X_x X_y - R^2}{R(X_x + X_y)} \tag{27}$$

and

$$e^{A\gamma} = \frac{\sqrt{(R^2 + X_x^2)(R^2 + X_y^2)}}{R(X_x - X_y)}. \tag{28}$$

Equation (27) can be used to confirm our previous choice of the location of the cut-off. At this frequency one of the two reactances, X_x and X_y, will be either resonant or anti-resonant. It is evident from (27) that if the phase shift is to have the desired value, $(n + 3/4)\pi$, at the assumed cut-off the non-resonant impedance must have the magnitude R. That this value is approximated follows from the symmetrical spacing of transfer constant and image impedance controlling frequencies with respect to the cut-off. On both sides of the transition interval, in the regions of uniform spacing of poles and zeros, the non-resonant reactance approximates $R \tan \pi f/2\alpha$ or $R \cot \pi f/2\alpha$ and at the middle of each space, where $\pi f/2\alpha$ is an odd multiple of $\pi/4$, is numerically equal to R. Hence, by symmetry, this must also be the value approximated at the middle of the non-uniform interval between the two chains, i.e., at the cut-off frequency.

Nature of the Approximation

The argument of Part I shows that the three-quarter spacing between the cut-off and the chain of transfer constant controlling factors results in poorer approximations to phase linearity in the transmission band and to complete suppression in the attenuating band than would the half-spaced cut-off solution. The three-quarter spacing between the cut-off and the chain of impedance controlling frequencies also leads to less perfect uniformity of the impedance characteristic. This is the price we pay for the larger range of phase linearity. Nevertheless, the error of approximation for both θ and Z_I if we follow the sense of equation (22) can be shown to be $\frac{1}{8}\alpha\left(\frac{1}{1-f}-\frac{1}{1+f}\right)$, when α is small, and hence can be made as small as we please by a suitable choice of α.[24] So far as the phase and impedance characteristics are concerned, experience shows that satisfactory precision can be obtained with a moderate value of α. The situation with respect to the attenuation characteristic is more serious. As we have already seen, the best approximation in the attenuating band is obtained by a cut-off spacing which is, if anything, slightly less than, rather than slightly greater than, $\alpha/2$. Furthermore, it appears from the above formula that with the three-quarter cut-off spacing, the approximation error at a given frequency in the attenuating band is proportional only to the first power of α. Hence cutting α in two, which substantially doubles the number of elements in the structure, adds but 6 db to the attenuation at this frequency. It is clear that a practical limit is thus set upon the suppression which can be provided.

Since the attenuation of the structure is relatively low, the contribution of reflection effects to the total loss is correspondingly important. A peak of loss occurs at each impedance controlling frequency, where the lattice impedances are zero or infinite together.

[24] It is not true that the error in $\partial\beta/\partial\omega$ vanishes with α. However, in the following example, which may be taken as typical, the variation of $\partial\beta/\partial\omega$ is still only about 1 per cent of its average value.

At these frequencies the image impedance changes sign, and therefore also the constant term of equation (25). Thus, although the phase slope is uniform throughout the attenuating range, the phase characteristic itself suffers discontinuities of π radians at each impedance controlling frequency. Whether the discontinuity is an increase or a decrease of π radians is not distinguishable for a non-dissipative network. When parasitic dissipation is taken into account the peaks are finite and the phase increases or decreases according as the line- or cross-arm of the lattice has the smaller resistance component at the peak frequency. The infinite peak at this frequency, and the associated abrupt change in phase, can evidently be restored by adding additional resistance to the smaller impedance so as to bring the arms into balance.

This observation is of importance in considering the effect of dissipation on the phase shift. A counterpart of Mayer's theorem can be found which relates the change in phase shift resulting from uniform dissipation in the network elements to the slope of the loss curve. The formula is

$$\Delta B \doteq - \omega d \frac{\partial A}{\partial \omega},$$

where d is the dissipation constant, and where A and B are in népers and radians respectively. In the transition interval, where the slope of the loss curve is great, the effect of uniform parasitic dissipation may reduce the phase appreciably. This effect can be compensated by small modifications in the theoretical frequency spacings, or by the introduction of a lumped resistance to balance the bridge at the first impedance controlling frequency, according to the plan suggested above.

Example

To illustrate the performance of this sort of network, we may consider a low-pass filter containing four evenly spaced critical frequencies in the practical transmission band. Subsequent natural frequencies will then occur at 4.75α, 5.5α, 6.5α, etc., according to the rule for three-quarter spacing adjacent to the cut-off. We may suppose that the requirement for linearity of phase shift does not extend above 7.5α, so that the sequence of uniformly spaced impedance controlling frequencies may be terminated after this point according to the scheme proposed in the case of the high-pass filter. In the frequency range of interest, we can replace the omitted chain of uniformly spaced frequencies by a single natural frequency at double spacing. The transfer constant and image impedance expressions can then be written as:

$$\tanh\frac{\theta}{2} = i\frac{\pi f}{2\alpha}\frac{\left(1-\frac{f^2}{(2\alpha)^2}\right)\left(1-\frac{f^2}{(4\alpha)^2}\right)}{\left(1-\frac{f^2}{\alpha^2}\right)\left(1-\frac{f^2}{(3\alpha)^2}\right)\sqrt{1-\frac{f^2}{(4.75\alpha)^2}}},$$

and

$$Z_I = R\frac{\sqrt{1-\frac{f^2}{(4.75\alpha)^2}}\left(1-\frac{f^2}{(6.5\alpha)^2}\right)\left(1-\frac{f^2}{(9.5\alpha)^2}\right)}{\left(1-\frac{f^2}{(5.5\alpha)^2}\right)\left(1-\frac{f^2}{(7.5\alpha)^2}\right)}.$$

These equations determine Z_x and Z_y, the values of which may be used in equations (27) and (28) to calculate the performance. This is shown by Fig. 20, after dissipative effects have been taken into account.

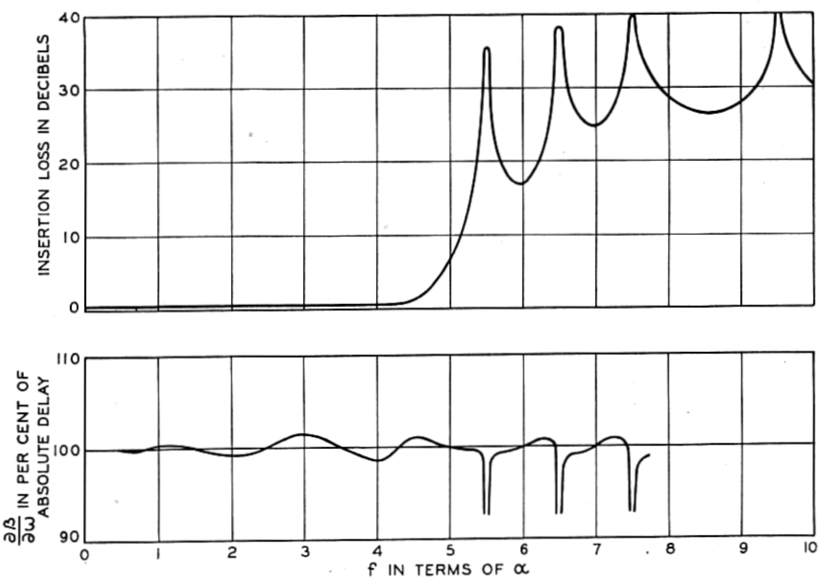


Fig. 20—Performance of a low pass filter having linear phase through the cut-off.

The approximation of the phase characteristic to linearity has again been indicated by exhibiting departures of the slope from the average.

It is observed that the approximation obtained by the three-quarter spacing is as close in the transition interval from 4α to 5.5α as in the practical transmitting band below 4α. In practical design problems, the phase shift is unlikely to be of interest beyond the first or second reflection peak, so that the chain of impedance controlling frequencies might be sooner terminated.

The loss characteristic reveals that no very high degree of suppression is attained. In fact, the loss falls to about 16 db in the trough beyond the first reflection peak. So serious a prejudice in favor of the phase characteristic would render the design unsuitable for certain engineering purposes. There are open, however, several possibilities for increasing the attenuation. Small modifications in the theoretical design parameters of the type which have been described, and in particular, slight separations of the theoretically coincident impedance controlling frequencies in the two arms of the lattice, enable the loss to be somewhat improved without much degradation of the phase characteristic. If very much higher attenuation is demanded, it can be provided by two simple structures of this type, separated by a resistance pad to preserve the reflection effects upon which the phase characteristic depends.

Further possibilities are suggested by combination of two principles already developed. It has been observed that a reduction in the three-quarter spacing of the cut-off would improve the selectivity of the structure but would also unduly increase the slope of the phase characteristic in the transition interval. We have also seen, however, that the result of uniform dissipation in the network elements is to diminish the phase shift in this region. Hence our analysis suggests that we may be able to obtain the desired phase characteristic in conjunction with the shorter cut-off spacing necessary for high selectivity if we deliberately increase the dissipation in the network.

A concomitant result of such procedure is seen to be an increase in the uniform loss in the transmission band, which may not always be desirable. Neither does the attempt to provide the phase property without sacrifice of high loss through the introduction of uniform dissipation represent the most effective attack on the problem. To achieve this end, resistances must be associated with the reactive elements of the lattice impedances in a precisely determined manner, not to be deduced solely from the foregoing theory of reactive networks. The elaboration of the theory to include also resistive impedance elements serves to determine a filter whose attenuation changes continuously from a low, uniform value in the pass-band to an arbitrary value in the attenuation bands with linearity of phase shift and, in addition, the third ideal property of constant impedance. The general theory, however, can more appropriately form the subject of a subsequent paper.

The solution of this problem completes the application of the methods for realizing ideal filter properties. We have seen that if all the properties are of importance and the desired approximations close, we are led to networks which, while formally simple, involve correspondingly numerous natural frequencies. On the other hand, if the impedance, or the phase, or the loss property be subordinated in respect to the others, suitable modification of the analysis allows the remaining properties to be realized with simplification of the structure.