Optimum Design of Multihole Directional Couplers With Arbitrary Aperture Spacing

Homayoon Oraizi, *Member, IEEE*

***Abstract*— A numerical procedure based on the method of least squares is presented for the optimum design of multihole directional couplers with a specified number of apertures and their arbitrary spacing in a desired frequency band for a given coupling coefficient (C) together with maximization of the mean value of isolation (I) or equivalently directivity (D) across the band and also specified values of (C) and (D) (or equivalently isolation I). An error function is constructed in the band for optimization of C and D or I, which after minimization gives the best values of the hole radii and spacing. The proposed method is quite suitable for computer-aided design (CAD) and has several advantages over common design methods for multihole directional couplers, such as simplicity of analysis and development of the numerical algorithm, flexibility of hole spacing, specification of the desired values of C and D or I, and actual optimization of the coupler design in the desired frequency band. The proposed method is potentially applicable to the design of other couplers.**

I. Introduction

DIRECTIONAL couplers are microwave passive four-port devices which have different applications in microwave networks and systems. Various types of couplers have been devised for waveguide, stripline, and microstrip circuits, but they are narrow-band devices which have the desired performance only in the close vicinity of the design frequency. Tandem connection of several couplers improves the frequency response, and its optimum design provides a broad-band device [1]–[10].

The Bethe-hole coupler is composed of two adjacent rectangular waveguides with a small hole in their common broad or narrow wall. To increase the bandwidth of this device, a train of holes is made in the common wall of the guide and is called a multihole directional coupler, as shown in Fig. 1. As the number of holes $(N+1)$ and the length of the hole train increase, the frequency response improves. The distance of the hole train from the guide edge(s) is chosen arbitrarily, and the hole spacing is kept constant and equal to a quarter-wavelength of the center frequency of the band $(d = \lambda_g/4)$. Therefore, the coupler design reduces to the determination of the hole radii. To achieve a binomial or maximally flat rippled frequency response of the directivity, coupling factors are made proportional to the binomial coefficients. For a Chebyshev response, the corresponding term in the expression of directivity is equated to the Chebyshev polynomial of degree N (for the number of holes of $N+1$) [1]–[6]. The Chebyshev coupler is almost optimum since, as we will see in this paper, the hole spacing of the optimum coupler is not quite equal to $\lambda_g/4$ of the center frequency of the band. A procedure for improving the rippled response of directivity is also provided [11], [12].

The coupler design based on the Chebyshev response is, in many cases, an overdesign, because the design of a broad-band coupler (with a large number of holes and $\lambda_g/4$ hole spacing) results in a relatively long device. In many applications, there is a space limitation and it is necessary to keep the device as short as possible. Furthermore, in many applications, an optimum possible directivity response is not required in the frequency band in the sense of minimizing the power leakage in the isolated port.

In this paper, a simple and effective numerical procedure based on the method of least squares is presented for the design of multihole directional couplers with arbitrary aperture spacing. An error function is constructed by comparing the desired coupling coefficient (C) with the coupled power and also the isolation (I) with the leakage power or maximizing the mean value of isolation (I) or equivalently directivity (D) across the desired frequency band. Minimization of the error function provides the hole radii, adjacent hole spacing (d_i), distance of each hole from the waveguide edge (s_i) for a given number of holes $(N+1)$ and a given C for the case where the mean value of D is to be as large as possible in the desired frequency bandwidth. In another procedure, the value of D or equivalently the value of $I = D + C$ is also initially specified. Then, the numerical procedures obtain the optimum hole spacing (d_i and s_i) and subsequently determine the new hole radii. Note that the adjacent hole spacings (d_i) and the distance of each hole from the guide edges (s_i) are not necessarily constant, as they are assumed so in other design methods. Furthermore, the proposed method is optimum in the least square sense, whereas the Chebyshev coupler is not strictly so, because the hole spacing is kept constantly equal to $\lambda_g/4$. Similar procedures have been applied for the design of stepline impedance transformers [17].

First, the numerical procedure is explained and then the results of computer programs are presented and compared with what is available in the literature.

II. Numerical Procedure

The top view of a multihole directional coupler, where the holes are made in the common broad wall of two adjacent rectangular waveguides, is shown in Fig. 1. The same figure may show the side view of a coupler with the common narrow wall. Since the field of the dominant TE_{10} mode is

Manuscript received January 22, 1997; revised January 14, 1998.
The author is with the Department of Electrical Engineering, Iran University of Science and Technology, Narmak, Tehran, 16844 Iran.
Publisher Item Identifier S 0018-9480(98)02752-5.

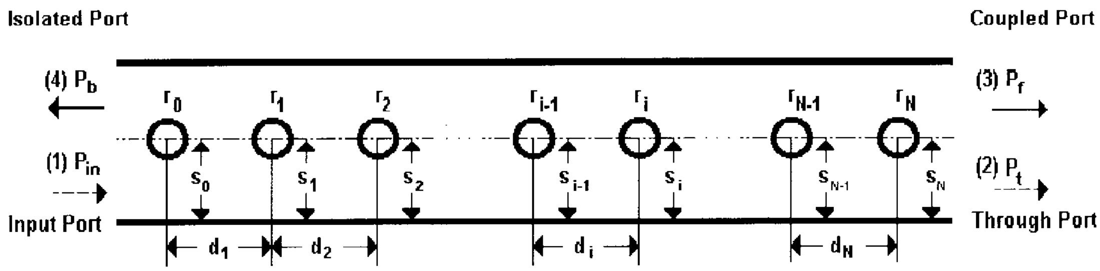


Fig. 1. Top or side view of a rectangular waveguide multihole directional coupler.

independent of the variable along the narrow wall, the distance of the holes in the common narrow wall from the guide edge(s) has no effect on the coupler performance. Referring to Fig. 1, the coupling coefficient (C), directivity (D), and isolation (I) are defined as follows:

$$C = 10\log_{10}(P_{\text{in}}/P_f) \tag{1}$$

$$D = 10\log_{10}(P_f/P_b) \tag{2}$$

$$I = 10\log_{10}(P_{\text{in}}/P_b) = C + D \tag{3}$$

where P_{in} is the incident power at the input port, P_f is the coupled power, and P_b is the power out of the isolated port [4].

The forward and backward traveling field amplitudes coupled through the nth hole from the input guide into the output guide at the kth frequency in the desired band (f_k) are, respectively [2], [13],

$$F_{nk} = K_{nk} r_n^3 = K_{nk} a_n \tag{4}$$

$$B_{nk} = L_{nk} b_{nk} \tag{5}$$

where r_n is the radius of the nth hole and $a_n = r_n^3$ and b_{nk} is defined below in (11). The coupling factors of the forward and backward traveling waves leaked through the nth hole in the common broad wall of the rectangular waveguide may be written as the following expressions [2, Sec. 8.4]:

$$\begin{aligned} K_{nk} &= \frac{A_{10}^{+}}{A} \\ &= j\frac{4\pi\sqrt{\varepsilon_r}f_k}{3abc\sqrt{1-(c/2a\sqrt{\varepsilon_r}f_k)^2}} \\ &\quad \cdot \left[\sin^2\left(\frac{\pi s_n}{a}\right) + 2\left(\frac{c}{2a\sqrt{\varepsilon_r}f_k}\right)^2 \cos\left(\frac{2\pi s_n}{a}\right)\right] \end{aligned} \tag{6}$$

$$\begin{aligned} L_{nk} &= \frac{A_{10}^{-}}{A} \\ &= j\frac{4\pi\sqrt{\varepsilon_r}f_k}{3abc\sqrt{1-(c/2a\sqrt{\varepsilon_r}f_k)^2}} \\ &\quad \cdot \left[2\left(\frac{c}{2a\sqrt{\varepsilon_r}f_k}\right)^2 - 3\sin\left(\frac{\pi s_n}{a}\right)\right] \end{aligned} \tag{7}$$

where s_n is the distance of the nth hole center from the guide edge, c is the speed of light in free space, a is the rectangular guide width, and ε_r is the dielectric constant of the insulating material inside the guide. The coupling factors of the forward and backward traveling waves through a hole in the common narrow wall of two adjacent rectangular waveguides are identical and equal to [1, p. 477]

$$\begin{aligned} K_k &= K_{nk} \\ &= L_{nk} \\ &= -j\frac{4\pi^2}{3b\beta_k}\left(\frac{1}{a}\right)^3 \\ &= -j\frac{4\pi}{3a^2b}[(2a\sqrt{\varepsilon_r}f_k\, c)^2 - 1]^{-1/2}. \end{aligned} \tag{8}$$

The total coupled fields traveling in the forward and backward directions are, respectively,

$$F_k = A\exp\left(-j\beta_k\sum_{i=1}^{N} d_i\right)\sum_{n=0}^{N} K_{nk}a_n \tag{9}$$

$$B_k = A\sum_{n=0}^{N} L_{nk}b_{nk} \tag{10}$$

where A is the field amplitude incident on the input port and, for simplicity, will be set equal to one $(A = 1)$, d_i is the distance between the $(i-1)$th and ith hole centers and

$$b_{nk} = a_n \exp\left(-j2\beta_k\sum_{i=1}^{n} d_i\right). \tag{11}$$

Note that mode interaction between closely spaced holes has been neglected.

According to the definitions of C, D, and I in (1)–(3), we set

$$C = -10\log_{10}\left|\frac{F}{A}\right|^2 \Rightarrow \left|\frac{F}{A}\right| = 10^{-(C/20)} = C_1 \tag{12}$$

$$I = -10\log_{10}\left|\frac{B}{A}\right|^2 \Rightarrow \left|\frac{B}{A}\right|^2 = 10^{-(I/10)} = I_1$$

$$D = -10\log_{10}\left|\frac{B}{F}\right|^2 \Rightarrow \left|\frac{B}{F}\right|^2 = 10^{-(D/10)} = D_1. \tag{13}$$

Now we would like to design a multihole directional coupler having $(N+1)$ holes, coupling coefficient (C) and directivity (D) (or equivalently, isolation $I = D + C$) in the frequency band from f_1 to f_u. We divide the desired frequency band into K discrete frequencies with the frequency interval equal to $\Delta f = (f_u - f_1)/(K-1)$. In the present method, the hole radii (r_i) and spacing (d_i and s_i) are determined starting from their specified initial values (as required by the version of the algorithm). Therefore, we construct an error function in the specified frequency band having the simplest mathematical form:

$$\varepsilon_1 = w\sum_{k=1}^{K}[-I_1 + |B_k|^2]^2 + \sum_{k=1}^{K}[-C_1 + |F_k|]^2$$
$$= w\sum_{k=1}^{K}\left[-I_1 + \left|\sum_{n=0}^{N} L_{nk} b_{nk}\right|^2\right]^2$$
$$+ \sum_{k=1}^{K}\left[-C_1 + \left|\sum_{n=0}^{N} K_{nk} a_n\right|\right]^2. \quad (14)$$

In case the coupling coefficient (C) is specified, but the mean value of directivity is to be maximized across the frequency band, the error function is constructed as follows:

$$\varepsilon_2 = w\sum_{k=1}^{K}|B_k|^2 + \sum_{k=1}^{K}[-C_1 + |F_k|]^2 \quad (15)$$

where w is a weighting factor which balances the contributions of the two terms in the error expression. Its value could be chosen in such a way as to enhance the effect of one term over the other. If one portion of the frequency band is more important than the rest, the weighting factor could be made a function of frequency to amplify the effect of that subband [11].

The error function for the design specification of C and D is constructed as follows:

$$\varepsilon_3 = w\sum_{k}\left[-D_1 + \left|\frac{B_k}{F_k}\right|^2\right]^2 + \sum_{k}[-C_1 + |F_k|]^2$$
$$= w\sum_{k}\left[-D_1 + \frac{|\sum_n L_{nk} b_{nk}|^2}{|\sum_n K_{nk} a_n|^2}\right]^2$$
$$+ \sum_{k}\left[-C_1 + \left|\sum_{n} K_{nk} a_n\right|\right]^2.$$

The algorithm based on ε_1 designs a coupler by simultaneous determination of the least mean squared errors of both the specified isolation as given by I_1 and the specified coupling as given by C_1 across the desired frequency band. That of ε_2 obtains the least output power from the isolated port (that is maximizing the mean value of isolation) together with the least mean squared error of C_1 across the band. That of ε_3 determines the least mean squared errors of both the specified directivity as given by D_1 and coupling as C_1 over the band.

Generally, the error function will have several minima with respect to the variables of hole radii and spacings. However, its absolute minimum point gives the optimum realization of the specified values of C and D. The minimization algorithms will at best locate a local minimum, which depend on the initial values of variables. We will use the gradient algorithms, which require the derivatives of the error function with respect to a_i, s_i, and d_i for $i = 0, 1, 2, \cdots, N$. (However, minimization routines without calculating derivatives could also be used [15], [16].) These operations can readily be performed for the error functions ε_1, ε_2, and ε_3, and are left out for brevity. The results may be specialized for the cases where $s_i = s$ and/or $d_i = d$.

To keep the algorithms simple, we assume that the hole train be along a straight line with a distance s from the guide edge, and that the distance between any two adjacent holes is equal to d. Therefore, for ε_1 we have

$$\varepsilon_1 = w\sum_{k=1}^{K}\left[-I_1 + |L_k|^2\left|\sum_{n=0}^{N} b_n\right|^2\right]^2$$
$$+ \sum_{k=1}^{K}\left[-C_1 + |K_k|\sum_{n=0}^{N} a_n\right]^2 \quad (16)$$

$$\frac{\partial\varepsilon_1}{\partial a_m} = 4w\sum_{k=1}^{K}\left[-I_1 + |L_k|^2\left|\sum_{n=0}^{N} b_n\right|^2\right]$$
$$\cdot |L_k|^2\sum_{n=0}^{N} a_n \cos[2\beta_k d(n-m)]$$
$$+ 2\sum_{k=1}^{K}\left[-C_1 + |K_k|\sum_{n=0}^{N} a_n\right]|K_k| \quad (17)$$

$$\frac{\partial\varepsilon_1}{\partial d} = 8w\sum_{k=1}^{K}\beta_k|L_k|^2\left[-I_1 + |L_k|^2\left|\sum_{n=0}^{N} b_n\right|^2\right]$$
$$\cdot \sum_{n=0}^{N-1}\sum_{m=n+1}^{N} a_n a_m (m-n)\sin[2\beta_k d(n-m)] \quad (18)$$

$$\frac{\partial\varepsilon_1}{\partial s} = 4w\sum_{k=1}^{K}\left[-I_1 + |L_k|^2\left|\sum_{n=0}^{N} b_n\right|^2\right]$$
$$\cdot |L_k|\frac{\partial}{\partial s}|L_k|\sum_{n=0}^{N}\sum_{m=0}^{N} a_n a_m \cos[2\beta_k d(m-n)]$$
$$+ 2\sum_{k=1}^{K}\left[-C_1 + |K_k|\sum_{n=0}^{N} a_n\right]\frac{\partial}{\partial s}|K_k|\sum_{n=0}^{N} a_n. \quad (19)$$

For ε_2 we have

$$\varepsilon_2 = w\sum_{k=1}^{K}|L_k|^2\left|\sum_{n=0}^{N} b_n\right|^2 + \sum_{k=1}^{K}\left[-C_1 + |K_k|\sum_{n=0}^{N} a_n\right]^2$$
$$= w\sum_{k=1}^{K}|L_k|^2\sum_{n=0}^{N}\sum_{m=0}^{N} a_n a_m \cos[2\beta_k d(n-m)]$$
$$+ \sum_{k=1}^{K}\left[-C_1 + |K_k|\sum_{n=0}^{N} a_n\right]^2 \quad (20)$$

$$\frac{\partial\varepsilon_2}{\partial a_m} = 2w\sum_{k=1}^{K}|L_k|^2\sum_{n=0}^{N} a_n \cos[2\beta_k d(n-m)]$$
$$+ 2\sum_{k=1}^{K}|K_k|\left[-C_1 + |K_k|\sum_{n=0}^{N} a_n\right] \quad (21)$$

$$\frac{\partial\varepsilon_2}{\partial d} = 2w\sum_{k=1}^{K}\beta_k|L_k|^2\sum_{n=0}^{N}\sum_{m=0}^{N} a_n a_m (m-n)$$
$$\cdot \sin[2\beta_k d(n-m)]$$
$$= 4w\sum_{k=1}^{K}\beta_k|L_k|^2\sum_{n=0}^{N-1}\sum_{m=n+1}^{N} a_n a_m (m-n)$$
$$\cdot \sin[2\beta_k d(n-m)] \quad (22)$$

$$\frac{\partial \varepsilon_2}{\partial s} = 2w\sum_{k=1}^{K}|L_k|\frac{\partial}{\partial s}|L_k|\left|\sum_{n=0}^{N}b_n\right|^2 + 2\sum_{k=1}^{K}\left[-C_1+|K_k|\sum_{n=0}^{N}a_n\right]\frac{\partial}{\partial s}|K_k|\sum_{n=0}^{N}a_n. \quad (23)$$

Since $\partial\varepsilon_2/\partial a_m$ is a linear function of a_m, the hole radii can readily be obtained by a mere matrix inversion as follows:

$$[a_n] = [1_{mn}]^{-1}[t_m] \quad (24)$$

$$1_{mn} = \sum_{k=1}^{K}\{w|L_k|^2\cos[2\beta_k d(n-m)] + |K_k|^2\} \quad (25)$$

$$t_m = C_1\sum_{k=1}^{K}|K_k|. \quad (26)$$

The coefficient matrix $[1_{mn}]$ and its inverse are symmetric. Their diagonal elements are also identical, as are the elements of the constant vector $[t_m]$. By writing the elements of the linear matrix equation for the various values of rows (m) and columns (n), it can be shown that for an even number of holes, the radii of the first half of the holes are one-to-one equal to those of the second half. For an odd number of holes, only the middle one is not paired. This situation also appears for the Chebyshev coupler.

The derivatives of the coupling factors with respect to s are

$$\frac{\partial}{\partial s}|K_k| = \mathrm{sign}(K_k)\frac{\partial}{\partial s}K_k \quad (27)$$

$$\frac{\partial}{\partial s}|L_k| = \mathrm{sign}(L_k)\frac{\partial}{\partial s}L_k \quad (28)$$

For ε_3, we have

$$\varepsilon_3 = w\sum_k\left[-D_1 + \left|\frac{L_k}{K_k}\right|^2\frac{|\sum_n b_n|^2}{(\sum_n a_n)^2}\right]^2 + \sum_k\left[-C_1+|K_k|\sum_n a_n\right]^2.$$

This algorithm is not pursued here.

The minimization of the error functions with respect to the variables a_n, s, and d are performed separately, so that the hole radii may be obtained for some specified values of s and d. The error function ε_2 is a quadratic function of a_m, and its minimum point with respect to them is obtained by a matrix inversion, although a gradient method could also be used as well. The method of interval halving is used for a linear search to obtain a local minimum along a direction, such as minimization of ε_1 and ε_2 with respect to s and d or along a gradient direction. However, the method of PARTAN is employed to minimize ε with respect to a_m [14]. This algorithm proceeds with alternate linear searches for the minimum point along the direction of the negative gradient and that constructed by connecting the minimum of the current gradient search and the starting point of the search along the preceding gradient direction. Initially, the first two linear searches are conducted along gradient vectors. Other gradient methods also proceed along directions other than the gradients, since the pure steepest descent method is not quite an efficient optimization algorithm, especially near singularities where it gets trapped in endless oscillations.

The method of interval halving is used for the linear search. The first step size along a direction vector is chosen equal to the distance between the initial point and the minimum of the quadratic surface approximating the error at that point:

$$DS = -\frac{2\varepsilon(\overline{Y})}{\langle \overline{P}, \nabla\varepsilon\rangle} \quad (29)$$

where $\varepsilon(\overline{Y})$ is the value of the error at the initial point $\overline{Y}$, $\nabla\varepsilon$ is the gradient vector at $\overline{Y}$, and $\overline{P}$ is the direction vector of the linear search. Therefore, the value of variables for this step size DS from their initial values along the direction vector $\overline{P}$ is

$$\overline{Y} \leftarrow \overline{Y} + DS\,\overline{P}. \quad (30)$$

However, the optimization should be performed under constraint, since the holes should not overlap each other and they should not get out of the guide. Therefore, the following conditions should be checked and appropriate corrections and modifications should be made after minimization of ε with respect to r_n, s, and d.

1) Enforce the following after minimization of ε with respect to r_n:

   if $r_n > r_{\min}$ continue
   else $r_n = r_{\min}$
   if $r_n < s$ and $r_n < a - s$ continue
   else $r_n = \min(s, a-s)$
   if $r_n + r_{n+1} < d$ continue
   else $r_n = \frac{r_n}{r_n + r_{n+1}}d$ and $r_{n+1} = \frac{r_{n+1}}{r_n+r_{n+1}}d$.

2) Enforce the following after minimization of ε with respect to s:

   if $s > 0$ and $s < a$ continue
   else $s = \max(r_n)$
   if $s > r_n$ and $a - s > r_n$ continue
   else if $s < a/2$ let $s = \max(r_n)$
   else $s = \min(a - r_n) = a - \max(r_n)$.

3) Enforce the following after minimization of ε with respect to d:

   if $d > \max(r_n + r_{n+1})$ continue
   else $d = \max(r_n + r_{n+1})$
   if $d < d_n$ continue
   else $d = d_u$
   if $r_n + r_{n+1} < d$ continue
   else $r_n = \frac{r_n}{r_n + r_{n+1}}$ and $r_{n+1} = r_{\frac{n+1}{r_n + rn+1}}d$.

where $r_{\min}$ is the minimum allowable hole radius and d_u is the maximum allowable length of the hole spacing.

Generally, the above algorithms fail for physically unrealizable couplers, where the specified length of the hole train is too short or the frequency bandwidth is too wide. In cases where the computer programs obtain a set of hole radii, which severely overlap each other due to the initial short length of the hole train, one could update the coupler length (L) after the initial determination of hole radii by the following formula and proceed thereon:

$$L = r_0 + r_N + 2\sum_{n=1}^{N-1}r_n. \quad (31)$$

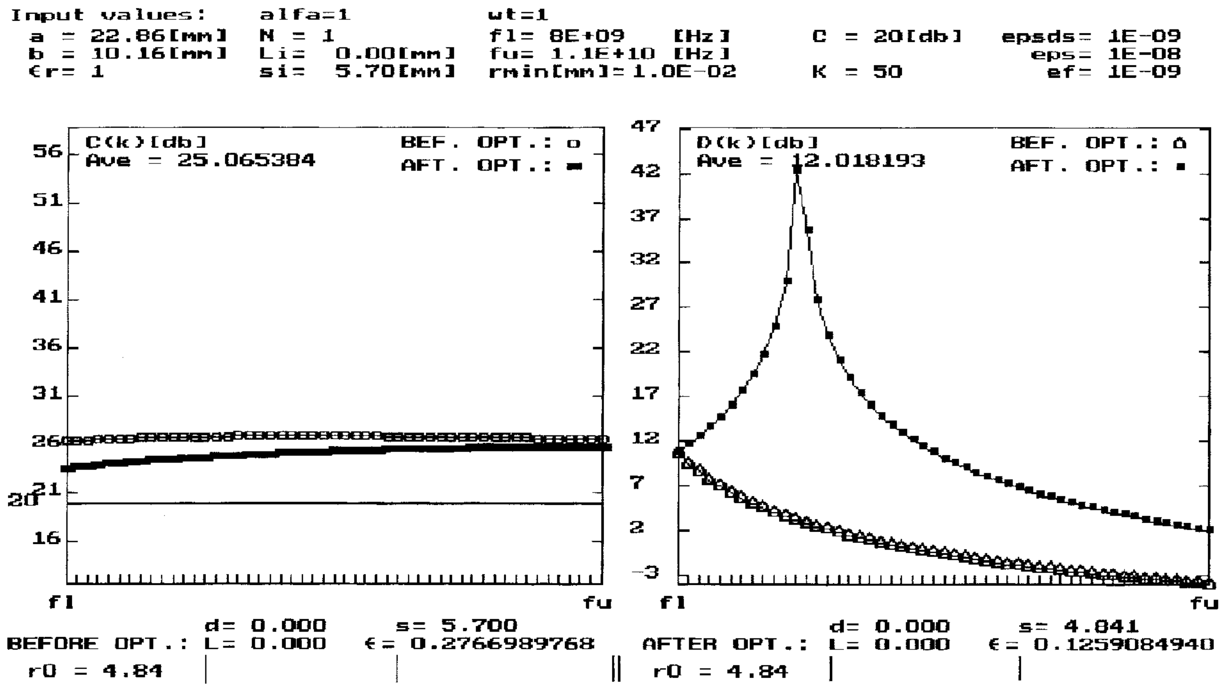


Fig. 2. A multihole directional coupler design with one hole in the common broad wall of two adjacent waveguides for specified coupling coefficient C.

The one hole coupler design formulas may readily be specialized from (16) to (26). For example, in the case that only the coupling coefficient (C) is specified (with the error function ε_2), the hole radius is obtained by (24)

$$a_0 = \frac{C_1 \sum_k |K_k|}{\sum_k [|K_k|^2 + w|L_k|^2]}$$

which pertains to the hole in the broad wall. Compare this result with [2, p. 404] where the design is based on a single frequency. For the hole in the narrow wall, we have

$$a_0 = \frac{C_1 \sum_k |K_k|}{\sum_k (1+w)|K_k|^2}.$$

III. Computer Programs and Numerical Results

Four computer programs have been written for the design of multihole directional couplers with holes in either the rectangular waveguide common narrow or broad wall and either the specified values of C (together with maximization of D) or specified values of C and D (or, equivalently, $I = C + D$). The inputs of the computer programs are: 1) rectangular waveguide cross-sectional dimensions (a and b); 2) dielectric constant of the material inside the waveguide (ε_r); 3) lower and upper frequency limits (f_1 and f_u) of the desired bandwidth (noting that the frequency bandwidth should be chosen so that only the dominant mode TE_{10} propagates); 4) number of discrete frequencies in the band (K); 5) number of holes (hereafter indicated by N and shown as such in the computer printouts); 6) initial specified length of the hole train $[L_i = (N-1)d_i]$; 7) initial specified hole spacing (d_i and s_i); 8) initial desired values of C and D (if required by the version of the computer program); 9) initial values of the hole radii; 10) maximum length of the hole train $(L_u = (N-1)d_u]$ where d_u is the maximum allowable distance between adjacent hole centers (say $d_u = \lambda_g/2$); 11) minimum value of hole radii $(r_{\min})$; and 12) stopping criteria to terminate the computer programs for the values of error (ef), error slope (epsds) and the least step size for interval halving (eps). The outputs of the computer programs are: 1) hole radii (r_n); 2) hole spacing (s and d); 3) length of the hole train (L); and 4) value of the error function (ε) with $w = 1$ before and after the optimization with respect to the hole spacing (s and d). The coupling (C) and directivity (D) response of the coupler with respect to frequency (namely C–f and D–f curves) are plotted before and after optimization.

Figs. 2–4 show 1-, 4-, and 20-hole directional coupler designs with holes in the broad wall of the rectangular waveguide for specified coupling coefficient ($C = 20$ dB) by the computer program based on the error function ε_2. Figs. 5 and 6 show four- and ten-hole coupler designs with the same specifications as above, and specifying C and isolation (I) by the program based on ε_1. Compare the results with [2, pp. 400–411].

Figs. 7–10 show 1-, 2-, 7-, and 20-hole coupler designs with holes in the narrow wall of waveguide for specified C by the program based on ε_2. Note that the one hole (in the narrow guide wall) coupler is actually a power divider. Figs. 11 and 12 show seven- and ten-hole coupler designs with the same specifications, and specified C and I by the program based on ε_1. Compare the results with [3, pp. 346–357]. The C–f and D–f curves are drawn before and after the optimizations. Note that the hole spacings of the optimum coupler designs are not equal to the quarter-wavelength ($\lambda_g/4 = 12.18$ cm) of the center frequency $f = 9$ GHz.

The following conclusions are arrived at by extensive computer program runs. The two algorithms based on the error function ε_2 are more effective for the coupler design than those based on ε_1, since ε_2 is a quadratic function of a_n and has a

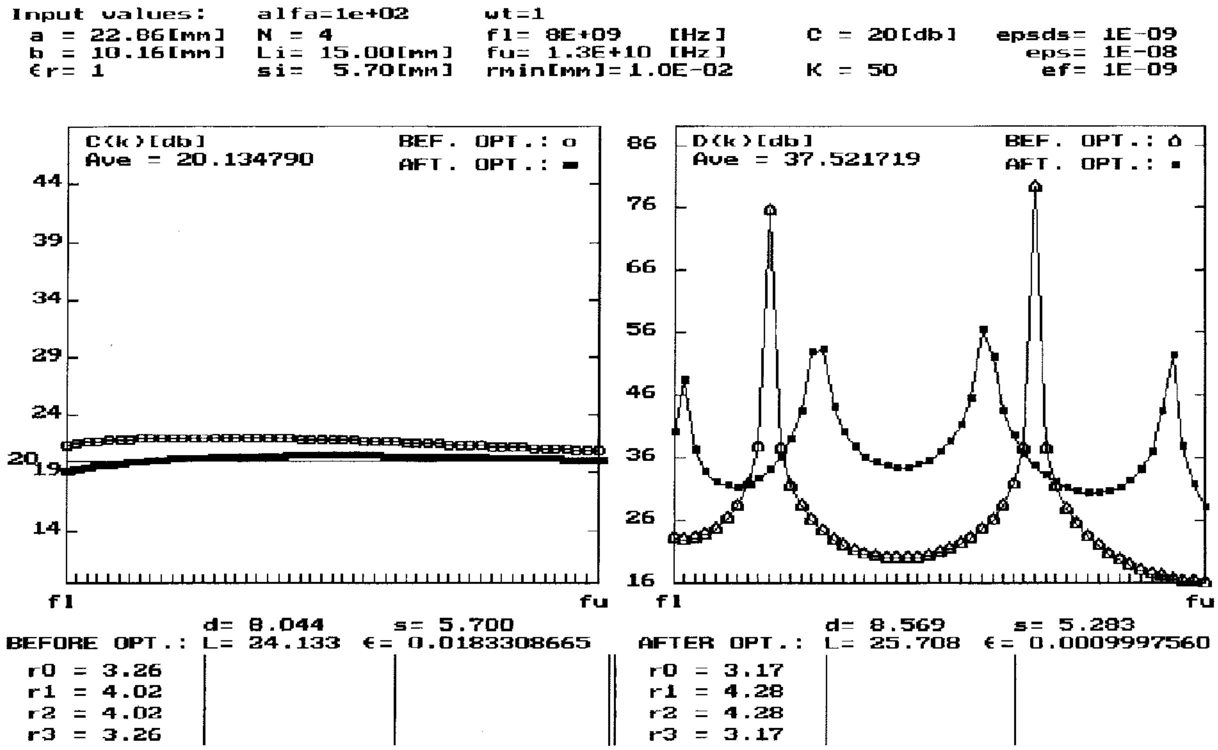


Fig. 3. A coupler design with four holes in the common broad wall for specified C.

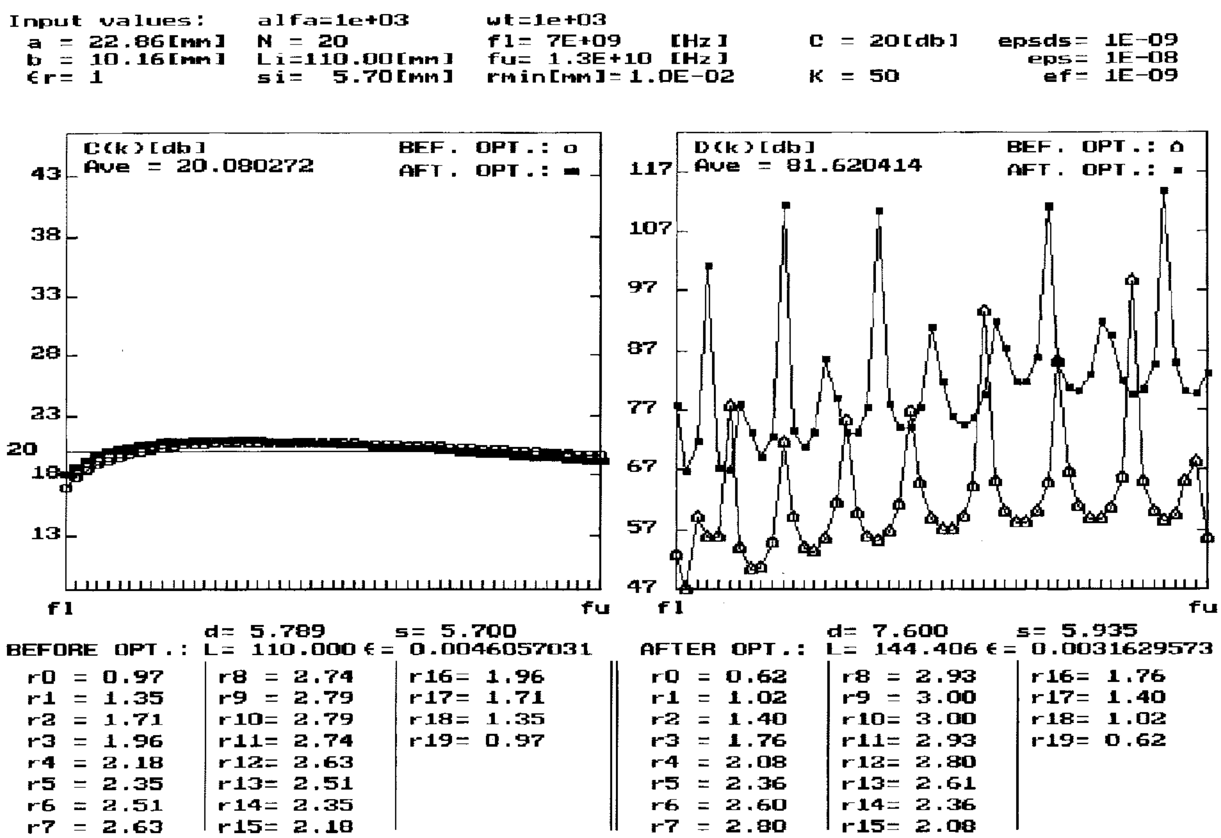


Fig. 4. Coupler C and D responses for 20 holes in the common broad wall with specified C.

single minimum with respect to a_n, though it is a complicated function of L and s. However, ε_1 is a nonquadratic function of a_n, L, and s, and requires specifying their initial values together with C and I. These algorithms obtain an optimum coupler design around the selected initial values for a_n, L, and s and are not expected to determine the best possible

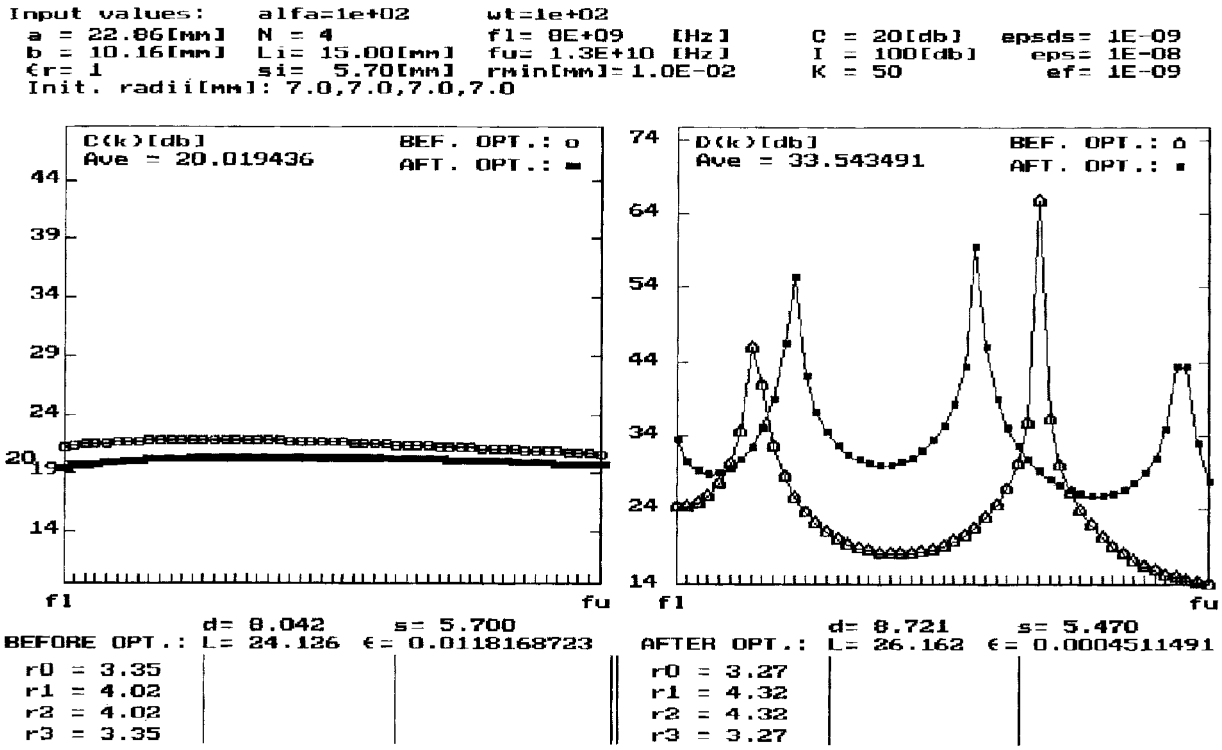


Fig. 5. Design of a coupler with four holes in the common broad wall for specified C and I.

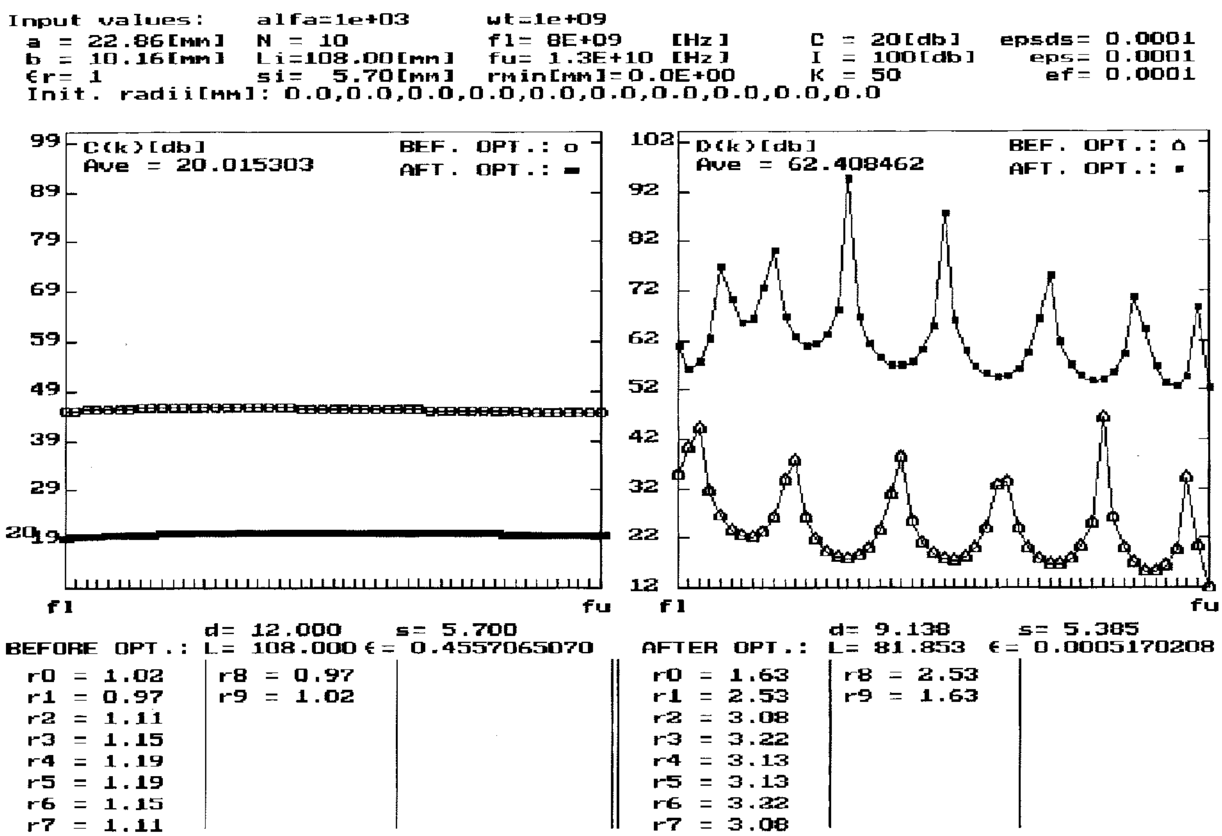


Fig. 6. Design of a coupler with ten holes in the common broad wall for specified C and I.

design; yet the algorithms based on ε_2, obtain the hole radii for any selected values of L and s and then optimize their values. However, the algorithms based on ε_1 obtain the best coupler for any combination of given values of a_n, L, and s, and then improve the design around these initial values, though the design is not necessarily strictly optimum. The

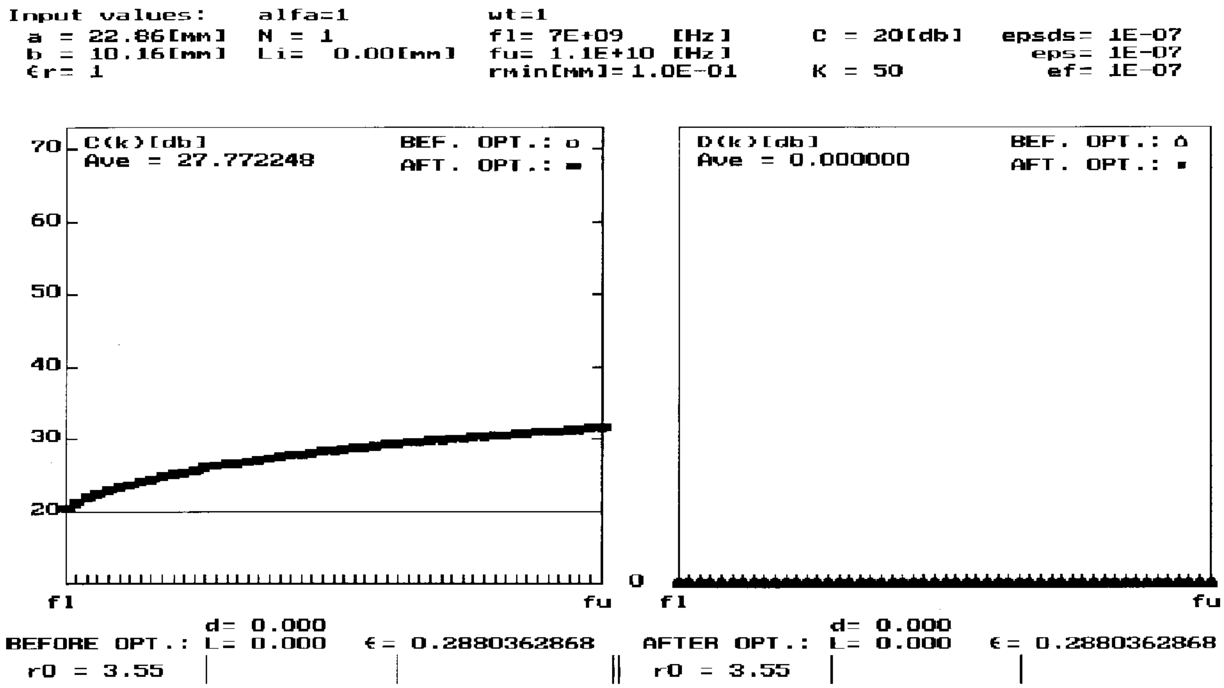


Fig. 7. Design of a power divider having one hole in the common narrow wall between two guides with specified C.

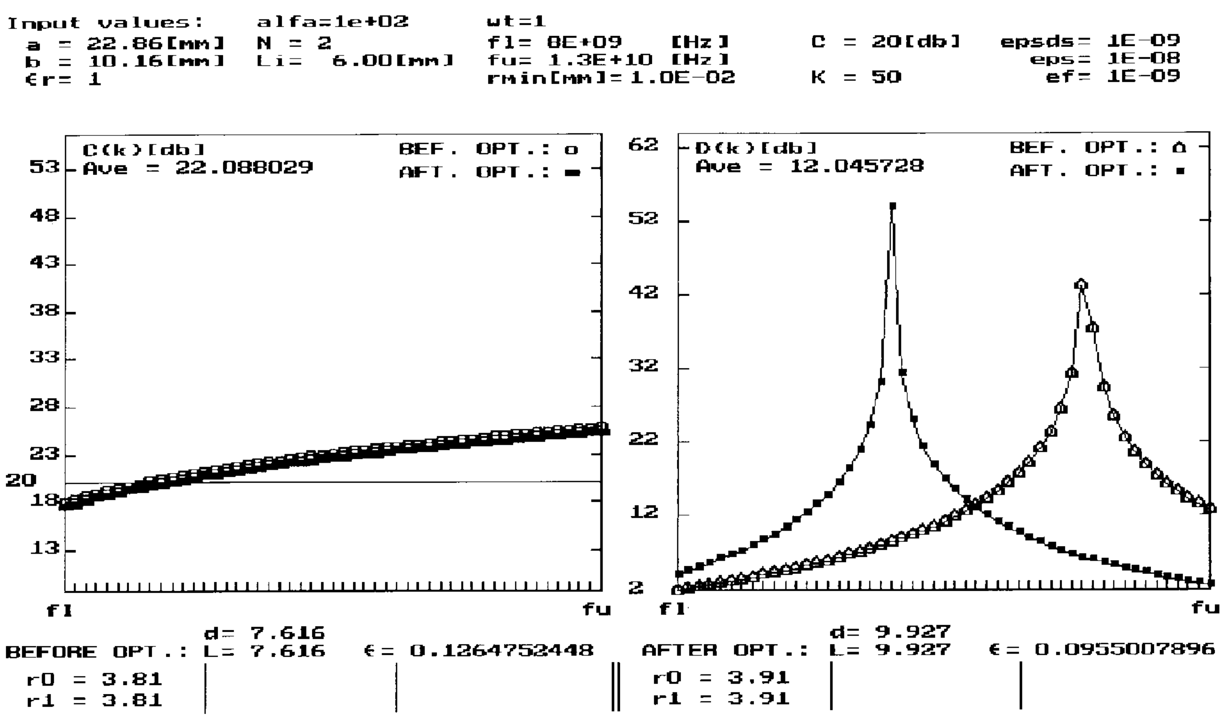


Fig. 8. Design of a coupler with two holes in the narrow wall for specified C.

algorithms based on ε_1 take much more central processing unit (CPU) time to converge toward the optimum coupler design than those based on ε_2. Increasing number of frequencies (K) and holes (N) increases CPU time, particularly for ε_1. Therefore, it may be advisable to start a coupler design with low values of K (say, 5) and N (say, 1) and then progressively increase N to the desired number of holes. This also gives an indication of the proper hole train length (L). Whenever the initial hole train length (L_i) is too short so that some holes overlap or the hole radii become too small or even negative and, consequently, the lower radius constraint is affected, then L_i should be increased accordingly. The optimum coupler design gives values for hole radii which start with a small value for the first hole increasing up to the middle hole and they symmetrically decreasing toward the last hole.

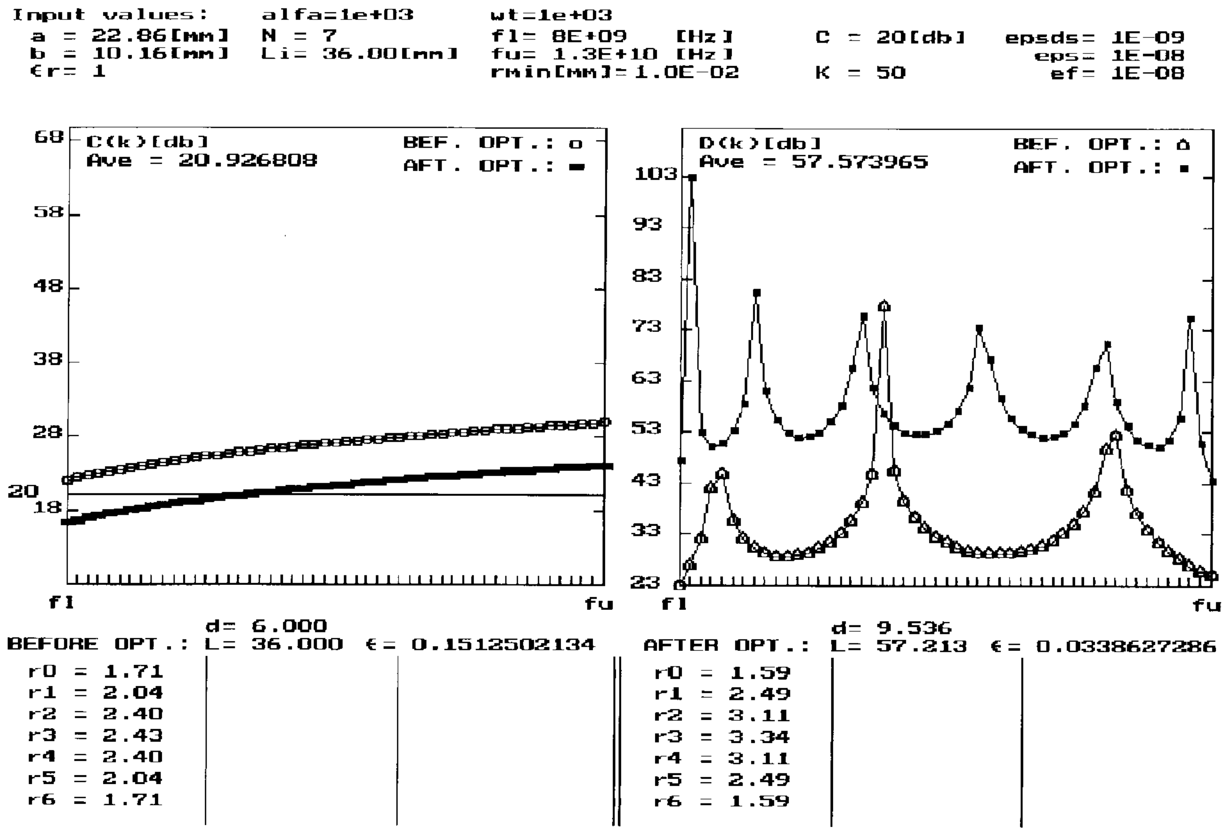


Fig. 9. A seven-hole coupler design with apertures in the narrow wall for specified C.

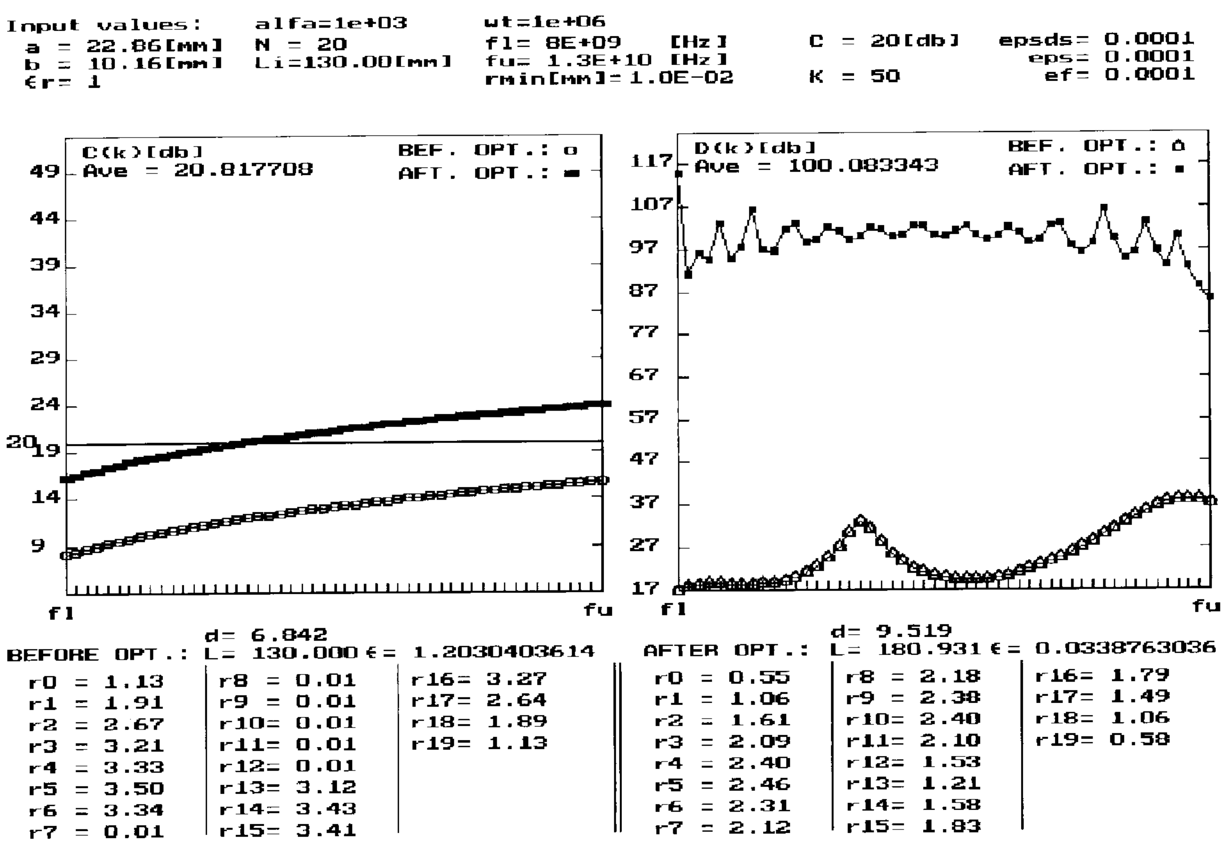


Fig. 10. A 20-hole-in-the-narrow-wall coupler with specified C.

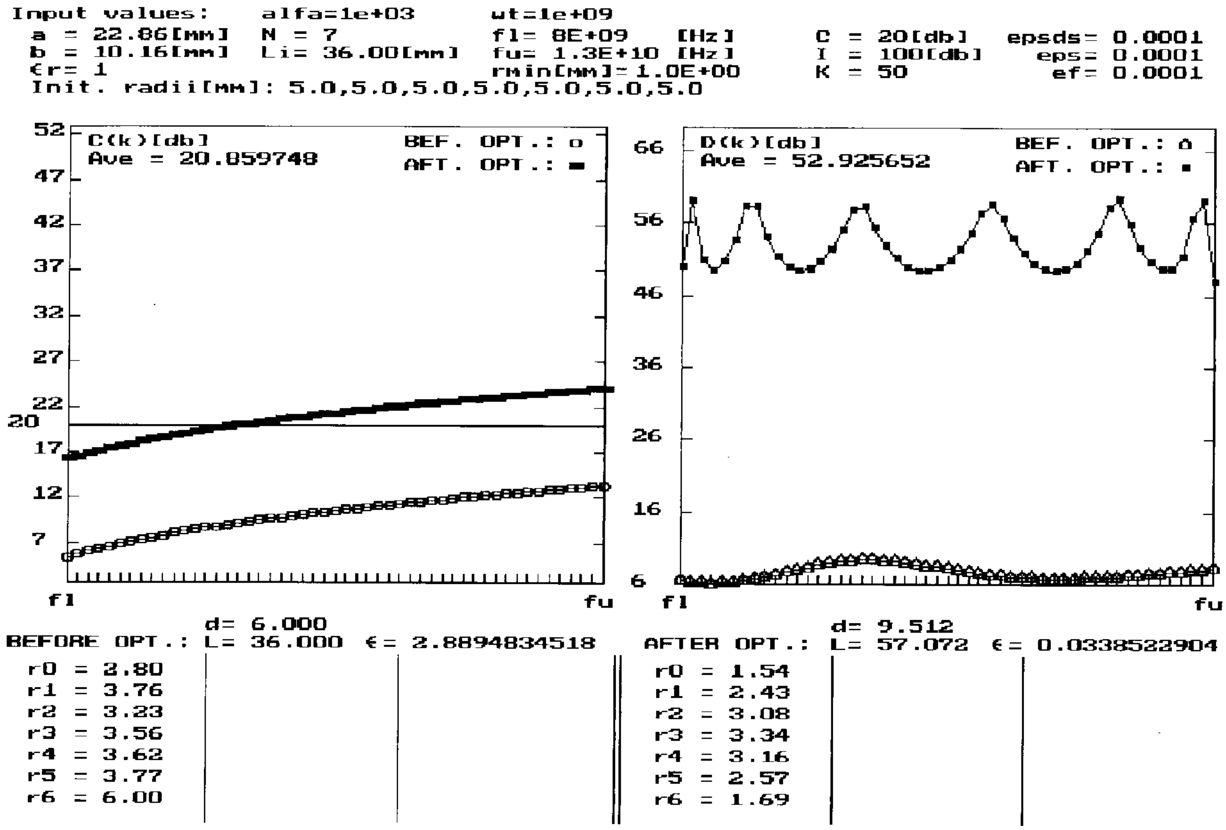


Fig. 11. A coupler with seven holes in the narrow wall for specified C and I.

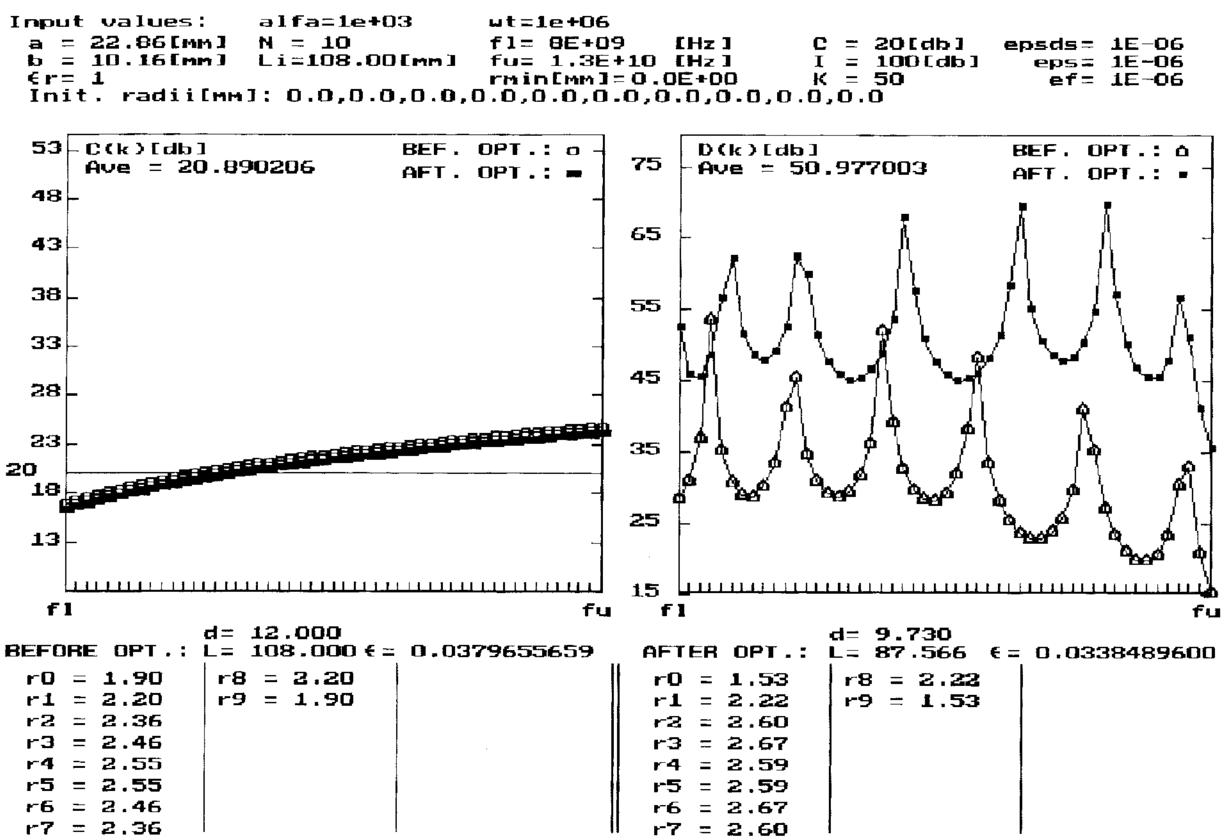


Fig. 12. A coupler design with ten holes in the narrow wall for specified C and I.

Better coupler designs should cause more ripples to appear in the D–f curve, since more ripples tend to improve the directivity response. The algorithms based on ε_1 perform best whenever convergence toward the optimum coupler design is quite slow and smooth, though CPU increases considerably. This is achieved by decreasing the interval halving stepsize (DS), which is implemented by dividing it by a factor α (alpha) and by decreasing the stopping criteria (epsds) up to 10^{-9} to 10^{-18}.

The value of stopping criteria (eps, epsds, and ef) have a definite effect on the convergence of algorithms, particularly the lower bound on interval halving as given by (eps) should be greater than about 10^{-9}, since our PC-486 hangs up.

The selected frequency bandwidth affects the convergence of the computer programs. Accordingly, the lower and upper frequency limits should not be close to the cutoff frequencies of the modes (say, TE_{10} and TE_{01}). Whenever the design value of coupling coefficient is achieved in the frequency band so that the C–f response is good, as for a large number of holes (N) and a long hole train length (L), then increasing the weighting factor (w) will considerably improve the D–f response. However, if the C–f response is not acceptable, then increasing w will worsen the C–f response. This procedure is quite effective for the algorithms base on ε_1, whereas the term related to D in the error function ε_2 is so constructed as to maximize the mean value of directivity (D), and the value of w has less effect on the convergence of the algorithm. Different values of w cause the D–f curve to shift with respect to the frequency axis, and eventually more ripples appear in its response, commensurate with the number of holes.

Large values of isolation (I) do not have much effect on the optimum coupler design and a good value for I is about 100. The algorithms based on ε_1 are quite sensitive to improper initial values of a_n, L_i, f_i, f_u, w, α, eps, and I. The algorithms design couplers with different combination of r_n, L, and s, which provide the specified coupling coefficient in the frequency band and strive to achieve the required directivity. Good designs are not usually unique, and small variations of r_n, L, and s considerably affect the C–f and, particularly, D–f responses.

It should be mentioned again that the programs first design the coupler for the specified initial values of hole spacings (s and d) and compute the hole radii and then minimize the error function with respect to s and d to improve and eventually optimize the coupler design. The examples show that the coupling coefficient (C) has little variation in the desired frequency band and its value is about the initially specified one after the first step of computation of the hole radii with the initial hole spacings (s and d).

Realization of the design values of the coupling coefficient and directivity is easier by the coupler with holes in the broad wall than one with the holes in the narrow wall, because the dominant TE_{10}-mode field is independent of the variable along the narrow wall. This is evident from the C–f curve of the broad-wall coupler, which has less variation in the frequency band and is nearer to the specified design value of C. Also note that the broad-wall coupler D–f curve has about one more ripple (for the same number of holes) than a narrow type, and the value of directivity in the first case is higher than in the latter.

If the computer programs are not able to design the coupler (i.e., computer errors occur), realization of the coupler is physically impossible with the specified value of d and it is required to increase the length of the hole train. In any case, optimization with respect to s and d substantially improves the directivity response and causes its value to increase and more ripples appear in the frequency band. The number of ripples in the D–f curve increases with increasing number of holes (N) and hole-train length (L). Realization of the desired C–f characteristics in the frequency band is generally simpler than the D–f characteristics. The former is nearly a horizontal curve in the band, whereas the variation of the latter is quite rippled. To improve the D–f response, it is necessary to increase the number of holes and the hole-train length.

IV. Conclusions

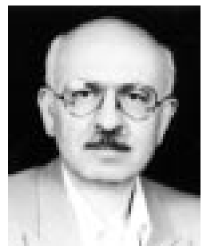
The algorithms developed here for the design of directional couplers based on the method of least squares, like any other optimization method, require specifying initial values as inputs to the computer programs. No procedure is developed here to give the best initial values for the hole radii, hole spacings $(d,\, s)$, weighting factor (w), step size for interval halving (DS/α), error criteria, specified value of isolation (I), lower and upper frequency limits of the band $(f_1,\, f_u)$, and number of frequencies. The problem of defining these initial values is quite involved, if at all tractable, and is really beyond the objectives of developing a numerical procedure for the coupler design. However, an error function should be constructed in such a way that it is as simple as possible and requires the least number of initial values. In this respect, the algorithm based on the error function ε_2 is superior to that of ε_1, since it does not require the initial values of hole radii. The computer programs may be written in an interactive form, so that the engineer may test the programs with various input values and finally arrive at an optimum coupler design. The comments made in the previous section may be useful. Any minimization algorithm at best locates a local minimum, which depends on the given initial values. The error function ε_2 is a quadratic function of the hole radii, but it is a nonquadratic function of L and s. The error function ε_1 is a complicated function of r_n, L, and s. On the other hand, with due regard to the above points, it is proposed that the algorithms developed here facilitate the design of multihole directional couplers with holes in the broad and narrow walls of the rectangular waveguides. They design a multihole coupler for some specified values of r_n, L, s, frequency bandwidth, C, and I. Subsequently, the programs optimize the initial designs. This could indeed be an advantage of these numerical procedures because they always given an answer, albeit nonoptimum, provided the coupler is physically realizable with the given specifications.

Acknowledgment

The author wishes to thank I. Rafiee for his assistance in writing the computer programs.

REFERENCES

[1] R. E. Collin, *Foundations for Microwave Engineering*, 2nd ed. New York: McGraw-Hill, 1992, ch. 6.
[2] D. E. Pozar, *Microwave Engineering.* Reading, MA: Addison-Wesley, 1990, ch. 8.
[3] R. S. Elliot, *An Introduction to Guided Waves and Microwave Circuits.* Englewood Cliffs, NJ: Prentice-Hall, 1993, ch. 10, 11.
[4] P. A. Rizzi, *Microwave Engineering: Passive Circuits.* Englewood Cliffs, NJ: Prentice-Hall, 1988, ch. 8.
[5] G. Matthaei, L. Young, and E. M. T. Jones, *Microwave Filters, Impedance-Matching Networks, and Coupling Structures.* Norwood, MA: Artech House, 1988.
[6] T. S. Saad, *Microwave Engineers' Handbook*, vol. 2. Norwood, MA: Artech House, 1971.
[7] W. E. Caswell and R. F. Schwartz, "The directional coupler—1966," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-15, pp. 120–123, Feb. 1967.
[8] R. Levy, "Directional coupler," in *Advances in Microwaves*, vol. 1, L. Young, Ed. New York: Academic, 1966, pp. 115–209.
[9] R. Levy, "Analysis and synthesis of waveguide multiaperture directional couplers," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-16, pp. 995–1006, Dec. 1968.
[10] R. Levy, "Improved single and multiaperture waveguide coupling theory, including explanation of mutual interaction," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-28, pp. 331–338, Apr. 1980.
[11] R. S. Elliot and Y. U. Kim, "Improved design of multihole directional couplers using an iterative technique," *IEEE Trans. Microwave Theory Tech.*, vol. 38, pp. 411–415, Apr. 1990.
[12] D. L. Swartzlander and R. S. Elliot, "Experimental study of multihole directional couplers providing a rippled response," *IEEE Trans. Microwave Theory Tech.*, vol. 40, pp. 1843–1846, Sept. 1992.
[13] R. E. Collin, *Field Theory of Guided Waves*, 2nd ed. Piscataway, NJ: IEEE Press, 1992.
[14] D. A. Pierre, *Optimization Theory and Applications.* New York: Wiley, 1969.
[15] M. J. D. Powel, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, vol. 7, pp. 155–162, 1964.
[16] J. W. Bandler, S. H. Chen, S. Daijavad, and K. Madson, "Efficient optimization with integrated gradient approximations," *IEEE Trans. Microwave Theory Tech.*, vol. 36, pp. 444–455, Feb. 1988.
[17] H. Oraizi, "Design of impedance transformers by the method of least squares," *IEEE Trans. Microwave Theory Tech.*, vol. 44, pp. 389–399, Mar. 1996.



Homayoon Oraizi (S'69–M'73) was born on April 24, 1942, in Isfahan, Iran. He received the B.E.E. degree from the American University of Beirut, Beirut, Lebanon, in 1967, and the M.Sc. and Ph.D. degrees in electrical engineering from Syracuse University, Syracuse, NY, in 1969 and 1973, respectively.

From 1973 to 1974, he taught at Khaje-Nassir University of Technology, Tehran, Iran. From 1974 to 1985, he worked for Iran Electronics Industries, Shiraz, Iran, as a Systems Engineer, Supervisor, and Head of the Systems Engineering Department, Communication Division, where he was engaged in various aspects of technology transfer, mainly in the field of HF/VHF/UHF communication systems. In 1985, he joined the Electrical Engineering Department, Iran University of Science and Technology, Narmak, Tehran, Iran, as an Assistant Professor. He is currently an Associate Professor in the Electrical Engineering Department, Iran University of Science and Technology. He teaches various electromagnetics engineering courses and has authored two books on electromagnetics and fields and waves (in Persian), has also translated an antenna book into Persian, and has published several papers in international journals and conferences. His research interests are in the area of numerical methods in electromagnetics engineering.