Improved coupled-mode theory for codirectionally and contradirectionally coupled waveguide arrays

R. R. A. Sym
d

Optical Devices Section, Department of Electrical Engineering, Imperial College, Exhibition Road, London SW7, UK

Received March 12, 1990; second revised manuscript received January 24, 1991; accepted January 24, 1991

A coupled-mode theory is derived from the scalar wave equation for the interaction between forward- and backward-traveling waves in a general slowly varying coupled waveguide array containing an arbitrary number of guides with a grating overlay. The equations include the effects of nonnegligible model overlap terms and changes in modal overlap but still demonstrate power conservation. Numerical results are presented for a contradirectional track-changing filter, which show that the modifications made to the equations are significant for practical devices.

1. INTRODUCTION

Grating devices have been extremely successful in guided-wave optics. They have been used in a wide range of applications, from passive filters to distributed Bragg reflectors in semiconductor lasers, and near-theoretical selectivity has been obtained in many cases. There are two underlying reasons. First, a waveguide can provide a well-defined wave direction and velocity, and, second, surface processing can permit the definition of a periodic structure with high accuracy.

Originally developed in integrated optics (see, e.g., Refs. 1 and 2), gratings have been demonstrated more recently in fiber optics. The most common configuration uses a grating overlay on half of a polished fiber coupler, but gratings have also been made directly on a D fiber. The use of a polished coupler has also permitted the fabrication of a track-changing reflector suggested in Refs. 9 and 10. Figure 1 shows the principle. First a polished coupler is made; different fibers are used in each half. One half uses a fiber that has a propagation constant \( \beta_A \), while the other uses a fiber with propagation constant \( \beta_B \). A grating is etched into one half, and the two are placed together in the normal way. If the grating wavelength \( \Lambda \) is such that \( K = 2 \pi / \Lambda = \beta_A + \beta_B \), the grating can phase match a forward wave in one fiber to a backward wave in the other. The result is a channel-dropping filter. A number of analogous transmission devices (which use a grating to phase match codirectional waves in asymmetric directional couplers) have also been proposed. Structures with still more coupled guides and a grating have also received limited attention. Most recently, a three-guide device has been shown to provide useful power-division functions, but arrays of arbitrary size have also been studied in distributed-feedback laser geometries. Almost all analyses have been performed with the use of the coupled-mode theory, and representative treatments of single-guide structures can be found in Refs. 17–20. The original study of a two-guide device was described in Refs. 9 and 10; this considered only two modes (one forward going, one backward) to be significant, but later theories included all four of the obvious modes (forward and backward waves in each guide). All subsequent coupled-mode models of more complicated grating-coupled arrays have been essentially similar.

The heart of any model involves a combination of the analysis of a directional coupler and the analysis of a grating. Recently the theory underlying the former—originally performed by Yariv and Taylor and Yariv—has been called into question. From the pioneering research of Hardy and Streifer, it has been clearly demonstrated that several of the assumptions and conclusions of the original models are incorrect, and a number of new theories have been derived. Because the revisions are complicated, I will devote space here to a brief review.

The conventional approach is to assume a solution as a sum of the modes of the individual guides. If there are \( N \) of them, the solution contains \( N \) modes. This solution is substituted into the wave equation, and some terms are eliminated with the use of the waveguide equations of the individual guides. Approximations are also made at this stage by assuming that the mode amplitudes vary slowly with distance. Next, an overlap integral is performed between the resulting equation and each of the individual waveguide fields in turn. This crucial process reduces one three-dimensional differential equation to a set of \( N \) one-dimensional ones that describe changes in the amplitudes of the modes in each guide. These equations contain coefficients in the form of overlap integrals, which typically couple the amplitude of any mode to all others.

Much of the inaccuracy of the old theory stems from its neglect of certain coefficients, despite retaining others. In particular, direct modal overlaps are often discarded, while overlaps involving modal fields and index perturbations (known as coupling coefficients) are kept. However, some coupling terms (self-coupling and non-nearest-neighbor coupling coefficients) are commonly ignored. It is also usual to assume that the small number of coefficients eventually retained have constant values, for example, independent of index perturbations owing to electro-optic switching. Finally, further inaccuracy follows from the incorrect definition of power that must be adopted for consistency with a neglect of direct overlaps.

These deficiencies are rectified in the new theories, recent reviews of which can be found in Refs. 27 and 28. Both vectorial and scalar versions have been developed.
The former include the Hardy–Streifer theory mentioned above and those of Haus et al. and Chuang. Attempts have been made to demonstrate equivalence among these three models, and, despite some controversy, it may be stated with reasonable confidence that all theories lead essentially to the same conclusions. Into the category of scalar theories fall those of Qian, Marcatili, and Peall and Syme. These are less accurate and valid only for weak guides. Nonetheless, they are scalar analogs of the different vectorial theories (see, e.g., Ref. 42 for a discussion of the precise correspondence) and reach similar conclusions within the restrictions of a TEM model.

All theories discussed so far are for uniform systems, whose parameters do not vary with distance. However, many devices have input–output transitions, in which the guides separate slowly. In the past these were modeled simply by a local adaptation of the uniform-array equations. A typical example of this adaptation is described in Ref. 44. Unfortunately, if one accepts that modal overlaps are nonnegligible, changes in overlap must also be significant. Further modifications are therefore needed for self-consistency. Two attempts have been made at a suitable vectorial theory, but in each case a rigorous derivation was not attempted, and the equations do not appear correct. Only one scalar model has been proposed, this has been derived from the theory of local normal modes and examined in some detail. The conclusion is that additional terms appear in the coupled-mode equations, which are required to satisfy power conservation, and account for redistribution of power as modal overlaps vary. Their effect is to modify device responses considerably. An unpublished vectorial theory of similar form has also recently been developed.

There is a considerable body of literature showing that the predictions of the old theory are approximately valid for a two-arm coupler. These references are far too numerous to cite here, but representative comparisons between experiment and theory can be found in Refs. 50–52 for Ti:LiNbO$_3$ devices. In light of the deficiencies outlined above, these results are somewhat surprising, so it is important to review the experimental evidence in favor of the revised theories. First, scalar modal overlaps of the form $<E_i,E_j>$ have been measured by Marcatili et al. Typical values are of the order of 0.1–0.2, clearly non-negligible compared with unity for $<E_i,E_i>$ and $<E_2,E_2>$. Second, the variation of $<E_i,E_j>$ with electric field has recently been measured in Ti:LiNbO$_3$ directional coupler geometries. This variation has also been found to be significant and in line with theoretical predictions. Third, asymmetry and finite extinction in the switch response of two-arm couplers in both Ti:LiNbO$_3$ (Ref. 53) and GaAs (Ref. 55) have been directly attributed to modal-overlap effects. Finally, the behavior of three-arm couplers has been shown to be vastly more complicated than was originally thought. In some cases, the error in the prediction of the old theory has been so large that 100% of the input power can emerge from the wrong guide.

Little has been done to incorporate the theoretical revisions described above into models of grating devices. The exceptions are analyses of a uniform system of two codirectionally coupled, grating-assisted guides (similar to those in Refs. 11–14), due to Marcuse and Huang. In the present paper I provide, for the first time to my knowledge, a unified theory of the new type for the interaction between codirectional and contradirectional waves in an array containing an arbitrary number of slowly varying single-mode guides combined with a reflection grating overlay. The analysis is scalar and based on a combination of previous models and the improved coupled-mode theory of Refs. 40–43. Though lossless operation is assumed, it is simple to incorporate loss or gain, and no major modifications to the equations are incurred in the process. It is therefore anticipated that a range of devices could be modeled in a similar way.

2. COUPLED-MODE THEORY

Derivation of the Equations

Figure 2 shows a grating-coupled waveguide array, formed from $N$ three-dimensional, single-mode guides lying roughly in the $z$ direction. The guides extend from $z = z_1$ to $z = z_2$, where they are decoupled. The grating extends from $z = 0$ to $z = d$, but, before tackling the complete problem, we shall briefly review a local normal-mode analysis of the individual guides. In the absence of the grating, each guide would be described in isolation by the scalar wave equation

$$\nabla^2 E(x, y, z) + n_i^2(x, y, z)k^2E(x, y, z) = 0 \quad (i = 1, 2, \ldots, N), \quad (1)$$

![Fig. 2. Geometry of general grating-coupled waveguide array.](image-url)
where \( E \) is the scalar electric field, \( \nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2 \), and \( k = 2\pi/\lambda \).

A solution could be attempted in the form of a sum of the local normal modes of the guide (i.e., the single guided mode plus a set of radiation modes), weighted by coefficients that vary with distance. The result would be a set of coupled differential equations showing how the modal coefficients change because of the structural variation of the guide. However, if the guide changes slowly compared with the beat length between any two such modes, there will be little mode conversion. In the limit of adiabatic variation, a solution can be assumed in the form of the single guided local normal mode. This varies locally, following

\[
E(x, y, z) = E_i(x, y, z)\exp\left[-j \int_0^z \beta(x)\,dz\right],
\]

where \( E_i(x, y, z) \) represents the local transverse field distribution of the guided mode and \( \beta(x) \) is its local propagation constant. The waveguide equation linking \( E_i \) and \( \beta \) is then given by

\[
\nabla^2 E_i + (n_i^2k^2 - \beta^2)E_i = 0,
\]

where \( \nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 \). Note that this implies that all bends are radiation free, and so on. We also assume that the fields are normalized, so \( \langle E_i(x, y, z) \rangle = \int_{\text{space}} E_i(x, y, z) \, dx\,dy = 1 \). We now consider the complete problem. The task is to solve a new scalar wave equation, given by

\[
\nabla^2 E(x, y, z) + n^2(x, y, z)k^2E = 0,
\]

where the refractive-index variation is

\[
n(x, y, z) = n_a(x, y, z) + \Delta n(x, y)\cos(Kz).
\]

Here \( n_a \) is the index distribution of the array (essentially an assembly of the \( n_i \) terms), again assumed slowly varying in the \( z \) direction, and \( \Delta n \) is a perturbation due to the grating. For simplicity, we take for the grating a constant spatial frequency \( K \), where \( K = 2\pi/\Lambda \) and \( \Lambda \) is the grating wavelength, and a \( z \)-independent index distribution \( \Delta n \).

A solution to Eq. (4) is now attempted in the form of a sum of forward- and backward-traveling local normal modes, namely, as

\[
E(x, y, z) = \sum_{i=1}^{N} E_i(x, y, z)\left[A_{Fi}(z)\exp\left[-j \int_0^z \beta_i(z)\,dz\right] + A_{Bi}(z)\exp\left[j \int_0^z \beta_i(z)\,dz\right]\right].
\]

Here \( A_{Fi} \) and \( A_{Bi} \) are the amplitudes of the forward- and backward-traveling modes, respectively, in the \( i \)th guide. Performing the relevant differentiation, substituting, eliminating terms by using Eq. (3) and neglecting slowly varying terms, we obtain, after some manipulation,

\[
\sum_{i=1}^{N} \left[ -2j\beta_i \partial E_i/\partial z \left( A_{Fi} \exp\left(-j \int_0^z \beta_i\,dz\right) - A_{Bi} \exp\left(j \int_0^z \beta_i\,dz\right) \right) \right] = 0.
\]

We now make the approximation that all propagation constants \( \beta_i \) are close to a reference value \( \beta_0 \), such that \( \beta_i(z) = \beta_0 + \Delta \beta_i(z) \) and \( K = 2\beta_0 \). Equating coefficients of terms of the form \( \pm j\beta_i(z) \) individually with zero and neglecting higher diffraction orders, we obtain two equations:

\[
\sum_{i=1}^{N} \left[ -2j\beta_0 \partial E_i/\partial z \left( A_{Fi} \exp\left(-j \int_0^z \Delta \beta_i\,dz\right) - A_{Bi} \exp\left(j \int_0^z \Delta \beta_i\,dz\right) \right) \right] = 0,
\]

\[
\sum_{i=1}^{N} \left[ 2j\beta_0 \partial E_i/\partial z \left( A_{Fi} \exp\left(-j \int_0^z \Delta \beta_i\,dz\right) + A_{Bi} \exp\left(j \int_0^z \Delta \beta_i\,dz\right) \right) \right] = 0.
\]

For convenience, the following substitutions are now made:

\[
F_i(z) = A_{Fi}(z)\exp\left[-j \int_0^z \Delta \beta_i(z)\,dz\right],
\]

\[
B_i(z) = A_{Bi}(z)\exp\left[j \int_0^z \Delta \beta_i(z)\,dz\right],
\]

which reduce Eqs. (8) to

\[
\sum_{i=1}^{N} \left[ E_i dF_i/dz + j\Delta \beta_i E_i F_i + j[k^2(n_a^2 - n_i^2)/2\beta_0] E_i B_i \right] = 0,
\]

\[
\sum_{i=1}^{N} \left[ E_i dB_i/dz - j\Delta \beta_i E_i B_i - j[k^2(n_a^2 - n_i^2)/2\beta_0] E_i F_i \right] = 0.
\]
The procedure is now to take inner products between Eqs. (10) and each modal field \( E_i \) in turn. The resulting set of \( 2N \) equations can be written as a set of two matrix-vector equations, in the form

\[
d\mathbf{F}/dz = [-j(\Delta \mathbf{\beta} + \mathbf{C}^{-1}\mathbf{K}^T) - \mathbf{C}^{-1}\mathbf{N}]\mathbf{F} - j\mathbf{C}^{-1}\mathbf{TB},
\]
\[
d\mathbf{B}/dz = [+j(\Delta \mathbf{\beta} + \mathbf{C}^{-1}\mathbf{K}^T) - \mathbf{C}^{-1}\mathbf{NB} + j\mathbf{C}^{-1}\mathbf{TF}],
\]

where \( \mathbf{F} \) and \( \mathbf{B} \) are \( N \)-element column vectors whose \( i \)th elements are \( E_i \) and \( B_i \), respectively, and \( \Delta \mathbf{\beta}, \mathbf{C}, \mathbf{K}, \mathbf{N} \), and \( \Gamma \) are \( N \times N \) matrices, whose \((i, j)\)th elements are given by

\[
\Delta \beta_{ij} = \Delta \beta \delta_{ij},
\]
\[
C_{ij} = \langle E_i, E_j \rangle,
\]
\[
K_{ij} = (k^2\beta_\perp)\langle E_i, (n_a^2 - n_\gamma^2)E_j \rangle,
\]
\[
N_{ij} = \langle E_i, \partial E_j/\partial z \rangle,
\]
\[
\Gamma_{ij} = (k^2/\beta_\perp)\langle E_i, n_a \Delta n E_j \rangle,
\]

\( \delta_{ij} \) is the Kronecker delta, and \( (A, B) = \int \int_\text{all space} A^* B \, dxdy \).

Equations (11) and (12) are the new system of equations that we set out to derive. We note that \( \mathbf{K}^T \) is a matrix describing codirectional coupling, in that its elements have the form of codirectional coupling coefficients. Likewise, \( \Gamma \) accounts for contradirectional coupling that is due to the grating, while \( \Delta \beta \) describes the degree of asynchronicity of the system. \( \mathbf{C} \) and \( \mathbf{N} \) account for the effects of direct modal overlap and of the power redistribution occurring when overlaps change, respectively. Both contain terms that were commonly neglected before, and which have subsequently been shown to be important. However, if we make the usual approximations and put \( \mathbf{C} = \mathbf{I} \) and \( \mathbf{N} = 0 \) (where \( \mathbf{I} \) and \( \mathbf{0} \) are the identity and zero matrices, respectively) the equations reduce to

\[
d\mathbf{F}/dz = -j(\Delta \mathbf{\beta} + \mathbf{K}^T)\mathbf{F} - j\mathbf{TB},
\]
\[
d\mathbf{B}/dz = +j(\Delta \mathbf{\beta} + \mathbf{K}^T)\mathbf{B} + j\mathbf{TF}.
\]

These equations have the form used in previous, less accurate models of track-changing reflective devices (e.g., Refs. 9, 21, and 22). We also note that, if there is no grating \( (\Gamma = 0) \), forward- and backward-traveling modes can be described separately by the system of equations

\[
d\mathbf{F}/dz = [-j(\Delta \mathbf{\beta} + \mathbf{C}^{-1}\mathbf{K}^T) - \mathbf{C}^{-1}\mathbf{N}]\mathbf{F},
\]
\[
d\mathbf{B}/dz = [+j(\Delta \mathbf{\beta} + \mathbf{C}^{-1}\mathbf{K}^T) - \mathbf{C}^{-1}\mathbf{NB}].
\]

These equations may be used to treat the grating-free transition regions at either end of the device. The upper one is equivalent to the existing scalar coupled-mode theory for slowly varying arrays.\(^{40-43}\) The lower one is what would be expected if the sign of \( z \) were to be reversed; note that this does not introduce a sign change in the term \(-\mathbf{C}^{-1}\mathbf{N}\) because of the \( z \) dependence of \( \mathbf{N} \). We conclude that the new theory is consistent with a wide range of previously published material.

**Boundary Conditions**

The device shown in Fig. 2 extends from \( z = z_1 \) to \( z = z_2 \) (at which point the guides are entirely decoupled), with a grating overlay from \( z = 0 \) to \( z = d \). Typically, operation may involve a forward-propagating input at \( z = z_1 \). The complete solution therefore involves the following steps:

1. Integration of the upper equation in Eqs. (14) from \( z = z_1 \) to \( z = 0 \), with boundary conditions of \( \mathbf{F}(z = z_1) = \mathbf{F}_{d1} \), to yield \( \mathbf{F}(z = 0) = \mathbf{F}_0 \).
2. Solution of Eqs. (11), with boundary conditions \( \mathbf{F}(z = 0) = \mathbf{F}_0 \) and \( \mathbf{B}(z = d) = \mathbf{0} \), to yield \( \mathbf{F}(z = d) = \mathbf{F}_d \) and \( \mathbf{B}(z = 0) = \mathbf{B}_0 \).
3. Integration of the upper equation in Eqs. (14) from \( z = d \) to \( z = z_2 \), with boundary conditions of \( \mathbf{F}(z = d) = \mathbf{F}_d \) to yield \( \mathbf{F}(z = z_2) = \mathbf{F}_{d2} \).
4. Integration of the lower equation in Eqs. (14) from \( z = 0 \) to \( z = z_1 \), with boundary conditions of \( \mathbf{B}(z = 0) = \mathbf{B}_0 \) to yield \( \mathbf{B}(z = z_1) = \mathbf{B}_{d1} \).

Of course, if the guides are sufficiently decoupled at \( z = 0 \) and \( z = d \), only step (2) is required.

**3. POWER CONSERVATION**

We can now prove that the new system of equations conserves power. According to conventional electromagnetic theory, the time-averaged flow of power in the \( z \) direction is given by

\[
P_z = \frac{1}{2\Re} \left( \int \int_\text{all space} (\mathbf{E} \times \mathbf{H}^*) \cdot \mathbf{k} dxdy \right),
\]

where \( \mathbf{E} \) and \( \mathbf{H} \) are time-independent vector representations of the electric and magnetic fields and \( \mathbf{k} \) is a unit vector in the \( z \) direction. If the normal scalar approximations are made, it is easy to show that for a solution in the form of Eq. (6), \( P_z \) can be reduced to

\[
P_z = (\beta_\perp/2\mu_0)\Re(\mathbf{F}^T \mathbf{C} \mathbf{F}^* - \mathbf{B}^T \mathbf{C} \mathbf{B}^*),
\]

where \( \omega \) is the angular frequency and \( \mu_0 \) the permeability. If power is conserved, then \( dP_z/dz = 0 \), which implies that

\[
(d\mathbf{F}^T/dz \mathbf{C} \mathbf{F}^* + \mathbf{F}^T \mathbf{dC}/dz \mathbf{F}^* + \mathbf{F}^T \mathbf{CdF}/dz) - (d\mathbf{B}^T/dz \mathbf{C} \mathbf{B}^* + \mathbf{B}^T \mathbf{dC}/dz \mathbf{B}^* + \mathbf{B}^T \mathbf{CdB}/dz) = 0.
\]

If, at this point, the coupled-mode equations are written in the alternative form,

\[
d\mathbf{F}/dz = \mathbf{M}_c \mathbf{F} + \mathbf{M}_d \mathbf{B},
\]
\[
d\mathbf{B}/dz = \mathbf{M}_c^* \mathbf{B} + \mathbf{M}_d^* \mathbf{F},
\]

where

\[
\mathbf{M}_c = -j(\Delta \mathbf{\beta} + \mathbf{C}^{-1}\mathbf{K}^T) - \mathbf{C}^{-1}\mathbf{N}, \quad \mathbf{M}_d = -j\mathbf{C}^{-1}\Gamma,
\]

then the power-conservation condition can be rewritten as

\[
\mathbf{F}^T (\mathbf{M}_c^T \mathbf{C} + \mathbf{M}_c^* + \mathbf{dC}/dz) \mathbf{F}^* - \mathbf{B}^T (\mathbf{M}_d^* \mathbf{C} + \mathbf{M}_d + \mathbf{dC}/dz) \mathbf{B}^* + \mathbf{F}^T (\mathbf{M}_d^* \mathbf{C} + \mathbf{M}_d) \mathbf{B}^* - \mathbf{B}^T (\mathbf{M}_c + \mathbf{M}_c^*) \mathbf{F}^* = 0.
\]

Equation (20) can be reduced to two simple conditions. From the first term in parentheses, we get

\[
\mathbf{M}_c^T \mathbf{C} + \mathbf{M}_c^* + \mathbf{dC}/dz = 0.
\]
The second term in parentheses gives the complex conjugate of this equation. Substituting the definition for $M_C$ given in Eqs. (19) and noting that $N + N^T = \frac{d}{dz}$, we find that Eq. (21) becomes

$$C\Delta \beta + K^T = \Delta \beta C + K. \quad (22)$$

Equation (22) implies that there must be a fixed relation among the coefficients of the matrices $C$, $\Delta \beta$, and $K$ of the form $K_1 - K_2 = (\Delta \beta_1 - \Delta \beta_2)C_0$. This was originally proved by Marcatili, who used Green’s theorem. Because the necessary mathematics is involved, we will not repeat it here; instead, we simply refer the reader to Ref. 39 or 42. Note that this relation implies that the new coupled-mode equations may be written in the altered form

$$\frac{dF}{dz} = [-jC^{-1}(\Delta \beta C + K) - C^{-1}N]F - jC^{-1}TB,$$
$$\frac{dB}{dz} = [+jC^{-1}(\Delta \beta C + K) - C^{-1}N]B + jC^{-1}TF. \quad (23)$$

Returning to Eq. (20), we see that the third term in parentheses requires that

$$CM_C^* - M_C^{*T}C = 0. \quad (24)$$

Again, the last term in parentheses gives the complex conjugate of this equation. Substitution of the definition of $M_C$ from Eqs. (19) yields $\Gamma = \Gamma^*$, which is self-evident. Power conservation is thus proved.

**4. NUMERICAL RESULTS**

No analytic solutions have yet been found to the system of equations. Here, we simply present numerical results, for the basic contradirectional track-changing filter geometry shown in Fig. 3. This consists of a pair of slab waveguides, with core indices $n_1$ and $n_2$ and widths $w_1$ and $w_2$, respectively. The substrate index is $n_s$ throughout. Each guide is bent through a large radius $r$, so that the interguide separation is a slowly varying parabolic function of $z$, with a minimum separation of $s$. The space between the guides contains a symmetrically placed sinusoidal grating, of period $\Lambda$, width $w_g$, and peak index change $\Delta n_p$. The grating is of length $L$, at which point the two guides are assumed to be so far apart that any modal overlap or contradirectional coupling is negligible. If the differential equations [Eqs. (11)] are integrated over the length $L$, subject to the boundary conditions previously outlined, the output powers from each guide are then given simply by $P_{21} = |F_1|^2$, $P_{22} = |F_2|^2$, $P_{11} = |B_1|^2$, and $P_{21} = |B_2|^2$. Similarly, the total output is the sum of these quantities.

Integration of the equations is a standard two-point boundary value problem, which can be solved with a Runge–Kutta algorithm. In this case, first the eigenvalue equation for each individual guide was solved numerically. The matrix coefficients were then evaluated at intervals along the device length. This was done by substituting the design data into analytic expressions previously found by carrying out the relevant overlap integrals analytically. Since the coefficients are all slowly varying, interpolation was used to calculate the values at intermediate positions.

The parameters were chosen to model a track-changing filter operation. This process revealed a number of interesting design problems. First, in such a device any codirectional coupling must be arranged to be asynchronous, which requires the two guides to have significantly different propagation constants. The important interactions—direct Bragg reflection and exchange Bragg reflection—can then be made synchronous individually, if the input wavelength is correctly chosen. However, in order roughly to equalize the coupling coefficients for these interactions, the grating must be narrow and strongly modulated. Otherwise, the direct Bragg coupling coefficients will be much larger than the exchange Bragg coefficient. Since there is a clear limit on the index change forming the grating (especially with weakly confirmed guides), the two guides must be close together for high reflectivity. The codirectional coupling coefficients then become extremely significant by comparison with the contradirectional coefficients, and the modal overlap terms also become large.

Figure 4 shows the variations of the codirectional coupling coefficients $K_{11}$, $K_{12}$, $K_{21}$, and $K_{22}$ along the length of the device for the following typical parameters: $n_s = 1.5$, $n_1 = 1.506$, $w_1 = 4 \mu m$, $n_2 = 1.505$, $w_2 = 4 \mu m$, $\Lambda = 0.5 \mu m$, $L = 1.5 cm$, $r = 1.6 m$, $s = 3 \mu m$, $w_g = 1.5 \mu m$, $\Delta n_p = 0.003$, and $\lambda = 1.503 \mu m$. The two coupling coefficients $K_{12}$ and $K_{21}$ are clearly different, a point that has been noted by many previous authors. Figure 5 shows similar variations of the contradirectional coupling coeffi...
Fig. 5. Variation of the contradirectional coupling coefficients $\Gamma_{11}$, $\Gamma_{12}$, and $\Gamma_{22}$ with distance along the device.

Fig. 6. Variation of the modal-overlap term $C_{12}$ and the power-redistribution term $N_{12}$ with distance.

coefficients $\Gamma_{11}$, $\Gamma_{12}$, and $\Gamma_{22}$, which have considerably smaller values than the codirectional terms. These have been roughly equalized through the choice of parameters. Finally, Fig. 6 shows the variations of the modal-overlap term $C_{12}$ and the power-redistribution term $N_{12}$. The peak value of $C_{12}$ is 0.317, which is clearly not small compared with unity. At the device ends (where the guide separation is $\approx 38 \text{ \mu m}$), all the coefficients are approximately zero.

All the coefficients above appear to be significant in determining the overall device response and show similar slow variations along the device length; however, they all show little change over the range of wavelengths of interest for filter operation. It was verified that the Marcatili relation $K_{12} - K_{21} = (\Delta \beta_1 - \Delta \beta_2)C_{12}$ was satisfied to high accuracy throughout.

On integration of the differential equations, it was found that power was conserved to an accuracy of $\approx 0.025\%$, with a moderate step size. Figure 7 shows numerical results for the device outputs as a function of optical wavelength, for an input to guide 1. The two large dips in the transmitted output $PF_1$ occur at $\lambda = 1.50263 \text{ \mu m}$ (exchange Bragg reflection) and $\lambda = 1.50316 \text{ \mu m}$ (direct Bragg reflection). The wavelengths concerned correspond only approximately to the values that are obtained from evaluation to the phase-matching conditions $K = \beta_1 + \beta_2 (\lambda = 1.50267 \text{ \mu m})$ and $K = 2\beta_1 (\lambda = 1.50303 \text{ \mu m})$, with the use of the values of $\beta_1$ and $\beta_2$ obtained by solution of the guide eigenvalue equations. This solution is due to the retention of all the terms in the matrix $C^{-1}K^T$ in Eqs. (11). Some of these terms have previously been neglected, but clearly they all have a significant effect. When the direct Bragg process occurs, power is coupled from the transmitted beam $PF_1$ into the backreflected beam $PB_1$. Similarly, when the exchange Bragg process occurs, power is coupled to the backreflected, cross-coupled beam $PB_2$.

The cross-coupled transmitted output $PF_2$ is negligible throughout. Despite this, examination of the mode amplitudes within the device showed that all four modes reach finite amplitudes at some point along the device length and must therefore be included in the calculation. This may explain why the reflectivity obtained in the direct Bragg process is higher than that in the exchange Bragg process, despite the fact that the coupling coefficients concerned are approximately equal.

For comparison, Fig. 8 shows numerical results for the same geometry, when modal overlaps are neglected (so

Fig. 7. Variation with wavelength of the output powers from the device, as predicted by the new coupled-mode theory.

Fig. 8. Variation with wavelength of the output powers from the device, assuming that $C = I$ and $N = 0$. 

R. R. A. Syms
ACKNOWLEDGMENT
It is a pleasure to acknowledge the stimulus provided by University. It is a pleasure to acknowledge the stimulus provided by including those terms previously neglected.

tical device response will therefore require the retention are in codirectional ones. An accurate prediction of practical device response will therefore require the retention in the differential equations of all the matrix coefficients, including those terms previously neglected.

REFERENCES