THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 51, No. 1, January, 1972
Printed in U.S.A.

Derivation of Coupled Power Equations

By D. MARCUSE

(Manuscript received August 6, 1971)

The modes of a multimode dielectric waveguide are coupled by imperfections of the waveguide structure. The propagation of the coupled modes is described by coupled wave equations involving the wave amplitudes. If the coupling functions are random variables the interaction between the modes can be described more easily by coupled power equations. The derivation of the coupled power equations from the coupled wave equations is accomplished with the help of perturbation theory.

I. INTRODUCTION

The interaction of the modes of a multimode waveguide can be described by coupled wave equations.[1,2,3] The coupling between the waves is caused by imperfections of the waveguide structure. These imperfections are either deviations of the refractive index from the index distribution of the perfect waveguide or they are departures of the waveguide geometry from its nominal value. Changes of the core diameter of an optical fiber causes coupling between the guided modes and also coupling of the guided modes to the radiation modes.

Solutions of the coupled wave equations are hard to obtain for many modes since not only the wave amplitudes but also their relative phases enter into the description.

In most problems of practical interest the coupling coefficient is a random function of distance and only the exchange of power between the modes is of interest. A description of this problem in terms of coupled wave equations yields more information (phase information) than is required and consequently is quite complicated. One might expect intuitively that a description in terms of power exchange between the modes should exist.[4] If it were permissible to add power instead of amplitude one would be tempted to write down power rate equations that account for the incremental loss of power of one mode in terms of the power that is transferred per unit length from this mode to all the other guided modes while an increase in power can be expressed by the power that is transferred per unit length from all the other modes to the mode under consideration. Coupled power equations for two modes have been derived from the coupled wave equations that admit just this intuitive kind of interpretation.[5] Our derivation of the coupled power equations from the coupled wave equations is not mathematically rigorous. It is based on ideas of perturbation theory and on the assumption that the coupling coefficients can be described by correlation functions with Gaussian shape. No attempt has been made to solve the difficult problem of estimating the accuracy of the approximate theory.

II. COUPLED WAVE EQUATIONS

Many problems of coupling between the modes of multimode waveguides or the coupling between several different transmission lines can be expressed in terms of coupled wave equations[1,2,3]

$$\frac{\partial a_\nu}{\partial z} = \sum_{\mu=1}^{N} c_{\nu\mu} a_\mu \,. \tag{1}$$

The amplitude of the νth mode is a_ν ; $c_{\nu\mu}$ is the coupling coefficient that describes the interaction between mode μ and ν. The diagonal elements of the matrix of coupling coefficients represent the propagation constants of the modes

$$c_{\nu\nu} = -i\beta_\nu \,. \tag{2}$$

The system of equations (1) is the starting point of our discussion. In the absence of loss, power must be conserved. We thus require

$$\frac{\partial}{\partial z} \sum_{\nu=1}^{N} | a_\nu |^2 = 0. \tag{3}$$

With the help of (1) we obtain

$$\frac{d}{dz} \sum_{\nu=1}^{N} a_\nu a_\nu^* = \sum_{\nu=1}^{N} \left(a_\nu \frac{da_\nu^*}{dz} + a_\nu^* \frac{da_\nu}{dz} \right) = \sum_{\nu=1}^{N} \sum_{\mu=1}^{N} (c_{\nu\mu}^* + c_{\mu\nu}) a_\nu a_\mu^* \,. \tag{4}$$

The asterisk indicates complex conjugation. In order to write the last part of (4) in its indicated form the ν and μ labels of the second term had to be interchanged. Since (4) must vanish for any possible choice of the amplitude coefficients (which at any point along the line can be chosen arbitrarily by means of initial conditions) we obtain the following condition for the coupling coefficients.

$$c_{\nu\mu}^* = -c_{\mu\nu} \,. \tag{5}$$

It is advantageous to separate the rapidly varying z-dependent part $\exp(-i\beta_\nu z)$ from the wave amplitudes by the relation

$$a_\nu = A_\nu e^{-i\beta_\nu z}. \tag{6}$$

The coupled wave equations now assume the form

$$\frac{dA_\nu}{dz} = \sum_{\mu \neq \nu} c_{\nu\mu} A_\mu e^{i\Delta\beta_{\nu\mu} z}, \tag{7}$$

with the abbreviation

$$\Delta\beta_{\nu\mu} = \beta_\nu - \beta_\mu . \tag{8}$$

It follows from this definition that

$$\Delta\beta_{\nu\sigma} + \Delta\beta_{\sigma\mu} = \Delta\beta_{\nu\mu} . \tag{9}$$

III. DERIVATION OF COUPLED POWER EQUATIONS

The coupling coefficients are functions of the z coordinate that measures distance along the waveguide axis. In metallic as well as dielectric waveguides the coupling coefficients assume the form

$$c_{\nu\mu} = K_{\nu\mu} f(z). \tag{10}$$

$K_{\nu\mu}$ is independent of z. The function $f(z)$ often describes the actual shape of the deformed waveguide boundary or the bent waveguide axis. From (5) we obtain the condition

$$K_{\nu\mu}^* = -K_{\mu\nu} . \tag{11}$$

The function $f(z)$ is supposed to be a stationary random variable whose correlation function is assumed to be Gaussian

$$\langle f(z) f(z-u) \rangle = \sigma^2 e^{-(u/D)^2}. \tag{12}$$

σ^2 is the variance and D is the correlation length of $f(z)$. Since we aim at deriving coupled power equations we use the fact that the average power carried by each mode is

$$P_\nu = \langle |a_\nu|^2 \rangle = \langle |A_\nu|^2 \rangle. \tag{13}$$

The symbol $\langle\ \rangle$ indicates an ensemble average. From (7) and (13) we obtain

$$\frac{dP_\nu}{dz} = \left\langle A_\nu^* \frac{dA_\nu}{dz} \right\rangle + \left\langle A_\nu \frac{dA_\nu^*}{dz} \right\rangle$$
$$= \sum_{\mu \neq \nu} \{ K_{\nu\mu} \langle f(z) A_\nu^* A_\mu \rangle e^{i\Delta\beta_{\nu\mu} z} + K_{\nu\mu}^* \langle f(z) A_\nu A_\mu^* \rangle e^{-i\Delta\beta_{\nu\mu} z} \}. \tag{14}$$

Let us assume that the waveguide begins at $z = 0$. At a point z sufficiently close to $z = 0$ we obtain the following approximate solution of the wave equations (7)

$$A_\nu(z) = A_\nu(0) + \sum_{\rho \neq \nu}^{N} A_\rho(0) K_{\nu\rho} \int_0^z f(x) e^{i\Delta\beta_{\nu\rho}x}\, dx. \tag{15}$$

The solution of the wave equations (15) is based on first-order perturbation theory. It applies if the coupling is sufficiently weak or if z is chosen sufficiently small so that $A_\rho(z)$ is only slightly different from $A_\rho(0)$. The input values $A_\rho(0)$ are not subject to statistical fluctuations; their value is thus identical with their average value. Neglecting terms of higher than second order in $K_{\nu\mu}$ we obtain by substitution of (15) into (14)

$$\frac{dP_\nu}{dz} = \sum_{\mu \neq \nu} \left\{ \sum_{\rho \neq \nu} K_{\nu\mu} K_{\nu\rho}^* A_\rho^*(0) A_\mu(0) \int_0^z \langle f(z) f(x) \rangle e^{i(\Delta\beta_{\nu\mu}z - \Delta\beta_{\nu\rho}x)}\, dx \right.$$
$$\left. + \sum_{\rho \neq \mu} K_{\nu\mu} K_{\mu\rho} A_\nu^*(0) A_\rho(0) \int_0^z \langle f(z) f(x) \rangle e^{i(\Delta\beta_{\nu\mu}z + \Delta\beta_{\mu\rho}x)}\, dx + \text{c.c.} \right\}. \tag{16}$$

Terms proportional to the first power of $K_{\nu\mu}$ are absent from (16) since

$$\langle f(z) \rangle = 0. \tag{17}$$

The abbreviation "c.c." indicates additional terms that are the complex conjugate of the terms shown in (16).

Using (12) and the relation

$$\Delta\beta_{\mu\rho} = -\Delta\beta_{\rho\mu}, \tag{18}$$

we can write

$$\int_0^z \langle f(z) f(x) \rangle e^{i(\Delta\beta_{\nu\mu}z + \Delta\beta_{\mu\rho}x)}\, dx = \sigma^2 e^{i(\Delta\beta_{\nu\mu} - \Delta\beta_{\rho\mu})z} \int_0^z e^{-(u/D)^2} e^{i\Delta\beta_{\rho\mu}u}\, du. \tag{19}$$

We must now assume that $z \gg D$ requiring simultaneously that σ^2 is sufficiently small so that (15) is a good approximation. We can then write

$$\int_0^z e^{-(u/D)^2} e^{i\Delta\beta_{\rho\mu}u}\, du = \int_0^\infty e^{-(u/D)^2} e^{i\Delta\beta_{\rho\mu}u}\, du$$
$$= \tfrac{1}{2}\sqrt{\pi}\, D e^{-[(D/2)\Delta\beta_{\rho\mu}]^2} + iF(D, \Delta\beta_{\rho\mu}). \tag{20}$$

The function $F(D, \Delta\beta_{\rho\mu})$ is real and independent of z. It is thus clear that the expression in (19) depends on z only through the factor $\exp\,[i(\beta_\nu - \beta_\rho)z]$. Oscillatory terms do not contribute appreciably to the right-hand side of the differential equation (16). This fact becomes apparent if we integrate (16) with respect to z. The functions $A_\nu(z)$ are slowly varying compared to the oscillatory terms. Integrals over sine and cosine functions contribute very little compared to integrals over slowly varying functions. This consideration allows us to neglect the oscillatory terms in favor of the nonoscillatory terms. We thus consider only the term (19) with $\beta_\nu = \beta_\rho$. In an exactly analogous manner we obtain also

$$\int_0^z \langle f(z)f(x)\rangle e^{i(\Delta\beta_{\nu\mu}z - \Delta\beta_{\nu\rho}x)}\,dx$$

$$= \sigma^2 e^{i(\Delta\beta_{\nu\mu} - \Delta\beta_{\nu\rho})z} \int_0^z e^{-(u/D)^2} e^{i\Delta\beta_{\nu\rho}u}\,du$$

$$= \sigma^2[\tfrac{1}{2}\sqrt{\pi}\, De^{-[(D/2)\Delta\beta_{\nu\rho}]^2} + iF(D, \Delta\beta_{\nu\rho})]e^{i(\beta_\rho - \beta_\mu)z}. \quad (21)$$

By the above argument, that only nonoscillatory terms contribute to the differential equation (16), we keep only the term with $\beta_\rho = \beta_\mu$. Equation (16) now assumes the form

$$\frac{dP_\nu(z)}{dz} = \sigma^2 \sum_{\mu \neq \nu} \Big\{ [K_{\nu\mu}K_{\nu\mu}^* A_\mu(0)A_\mu^*(0) + K_{\nu\mu}K_{\mu\nu}A_\nu(0)A_\nu^*(0)]$$

$$\cdot \left[\frac{\sqrt{\pi}}{2} De^{-[(D/2)(\beta_\nu - \beta_\mu)]^2} + iF(D, \Delta\beta_{\nu\mu})\right] + \text{c.c.}\Big\}. \quad (22)$$

Because of our assumption that the coupling is sufficiently weak so that the wave amplitudes do not change very much over the distance z we replace $A_\mu(0)$ with $A_\mu(z)$ and obtain from (22) with the help of (11) and (13)

$$\frac{dP_\nu(z)}{dz} = \sqrt{\pi}\,\sigma^2 D \sum_{\mu=1}^{N} |K_{\nu\mu}|^2 e^{-[(D/2)(\beta_\nu - \beta_\mu)]^2}(P_\mu(z) - P_\nu(z)). \quad (23)$$

The complex conjugate terms cause the imaginary part of (22) to disappear.

Our derivation of equation (23) provides us with the derivative of the power of the ν-th mode essentially at $z = 0$. However, we could have followed the same procedure by assuming that the wave amplitudes or the power carried by each mode was given at some point z' and could have calculated the power derivative at some adjacent point

$z' + \Delta z$. With the assumption of weak coupling we would then have obtained the derivative of the power of the νth mode at an arbitrary point z' along the waveguide in terms of the power in all the other modes at (or near) that same point. In a strict sense, the procedure just described contains the flaw that the wave amplitudes $A_\nu(z')$ are statistical quantities. At $z = 0$ we used the fact that the wave amplitudes could be arbitrarily but definitely prescribed to take $A_\nu(0)$ out of the ensemble average. This cannot be done if we replace $A_\nu(0)$ with $A_\nu(z')$. However, we can use the following device to escape from this dilemma. Instead of considering the full ensemble of waveguides we select subensembles in such a way that within each subensemble the wave amplitudes $A_\nu(z')$ have the same complex value (within a certain prescribed narrow range). Our derivation of (23) then holds for each member of the subensemble. By regrouping the full ensemble into subensembles in such a way that the amplitudes $A_\nu(z')$ within each subensemble are all very nearly identical we find that equation (23) holds for each of the subensembles. We now average (23) over the subensembles and obtain an equation, equal in form to (23), that applies to the average power of the full ensemble. (See Appendix.)

For a given value of the rms deviation σ, (23) is obviously more accurate for smaller values of the correlation length D. For $D \to 0$ the length of guide needed for an approximate solution of the wave equations shrinks to zero and the only approximation left in our derivation consists in neglecting the oscillatory terms in (16). For finite values of D we must impose a restriction on the allowed values of σ. Let K^2 and P indicate the maximum values of $|K_{\nu\mu}|^2$ and $|P_\mu - P_\nu|$ while $\Delta\beta$ indicates the smallest value of $\beta_\nu - \beta_\mu$. Equation (23) yields the inequality

$$\frac{dP_\nu}{dz} \leqq \sqrt{\pi}\, \sigma^2 DNK^2 P e^{-[(D/2)\Delta\beta]^2}. \tag{24}$$

N is the number of guided modes in the waveguide. According to our approximate derivation we had to require that P_ν change only slightly over a distance in the order of the correlation length D. We integrate (24) over a distance κD with κ being a number close to but larger than unity and obtain

$$\frac{P_\nu(z + \kappa D) - P_\nu(z)}{P} \leqq \kappa \sqrt{\pi}\, \sigma^2 D^2 N K^2 e^{-[(D/2)\Delta\beta]^2}. \tag{25}$$

Since the right-hand side of (25) must be much smaller than unity in order for our approximate derivation of (23) to be valid we obtain the following inequality for the square of the rms deviation

$$\sigma^2 \ll \frac{e^{[(D/2)\Delta\beta]^2}}{\sqrt{\pi}\,\kappa D^2 N K^2}. \tag{26}$$

We have found the interesting result that the approximate coupled power equations hold for large values of σ not only in the limit of very small D but also in the limit of very large D. The minimum value of the right-hand side of (26) as a function of D is obtained for

$$D = \frac{2}{\Delta\beta}, \tag{27}$$

which is the least favorable correlation length in terms of the validity of the coupled power equations. The inequality (26) with D given by (27) becomes

$$\sigma^2 \ll \frac{e(\Delta\beta)^2}{4\sqrt{\pi}\,\kappa N K^2}. \tag{28}$$

If σ^2 remains below the value given by the inequality (28) the coupled power equations hold for all values of D.

We have assumed so far that the waveguide modes suffer no loss. This assumption is an idealization that can easily be extended to the more general case. The waveguide modes do not only suffer heat losses but also lose power by radiation. In fact the same mechanism that causes coupling between the guided modes also causes coupling to the radiation field so that losses are inevitable if the modes are coupled. If mode ν suffers the power loss α_ν we obtain the coupled power equations in the presence of loss in the obvious form†

$$\frac{dP_\nu}{dz} = -\alpha_\nu P_\nu + \sqrt{\pi}\,\sigma^2 D \sum_{\mu=1}^{N} |K_{\nu\mu}|^2 e^{-[(D/2)(\beta_\nu-\beta_\mu)]^2}(P_\mu - P_\nu). \tag{29}$$

Equation (29) has a simple intuitive interpretation. The decrease of power per unit length in mode ν is caused by the loss of power $\alpha_\nu P_\nu$ per unit length to heat and radiation, by an outflow of power to all the other modes indicated by the index μ, and finally there is an influx of power from all the other modes to mode ν that tends to offset the power loss.

IV. CONCLUSIONS

For randomly coupled waves the coupled wave equations can be transformed into coupled power equations. In this paper we derived

† The assumption of a Gaussian correlation function (12) is not essential for our derivation. In general $\sqrt{\pi}\,\sigma^2 D \exp\{-[(D/2)(\beta_\nu - \beta_\mu)]^2\}$ can be replaced by the Fourier transform of the correlation function $\langle f(z)f(z-u)\rangle$.

the coupled power equations for the special case of coupling coefficients that can be expressed as the product of a constant term times a function $f(z)$ that is independent of the mode labels. We further assumed that the correlation function of the coupling coefficients have Gaussian shape. The derivation of the coupled power equations from the coupled wave equations is based on perturbation theory. The coupled power equations are thus only approximately valid. An inequality was derived that limits the allowable values of the rms deviation σ of the coupling function $f(z)$. The coupled power equations are easier to use than the coupled wave equations since the coefficient matrix of the coupled power equations is constant, real, and symmetric so that the theory of symmetric matrices can be used for their solution.

The accuracy of our approximate derivation of the coupled power equations is not known.

APPENDIX

It is the purpose of the argument involving the subensembles to prove that we can use in (16)

$$\langle A_\rho^*(z')A_\mu(z')f(z' + \Delta z)f(z' + z + \Delta z)\rangle$$
$$= \langle A_\rho^*(z')A_\mu(z')\rangle\langle f(z' + \Delta z)f(z' + z + \Delta z)\rangle. \tag{30}$$

We abbreviate the product of the wave amplitudes by x and the product of f with itself by y and write (30) in the simplified form

$$\langle xy\rangle = \langle x\rangle\langle y\rangle. \tag{31}$$

To show that (31) holds we introduce the probability W_i of finding any member of the full ensemble, the probability P_j of finding a particular subensemble and finally the conditional probability $p_{i,j}$ of finding the member i of the full ensemble in the subensemble j. These three probabilities are related by the equation

$$W_i = p_{i,j}P_j\,. \tag{32}$$

We can now write the ensemble average in the form

$$\langle xy\rangle = \sum_i W_i x_i y_i = \sum_i p_{i,j}P_j x_i y_i\,. \tag{33}$$

The sum over the full ensemble can now be rearranged to be extended first over a particular subensemble (this sum is indicated by a prime on the summation sign) followed by a sum over all the subensembles

$$\langle xy \rangle = \sum_j P_j \sum_i{}' p_{i,j} x_i y_i \ . \tag{34}$$

The subensembles were selected on the basis that x_i was constant (or nearly so) in any given one of them. We can thus take x_i out of the sum over the members of each subensemble and obtain

$$\langle xy \rangle = \sum_j x_j P_j \langle y \rangle_j \ . \tag{35}$$

The symbol $\langle \ \rangle_j$ indicates an ensemble average over the members of the jth subensemble. The subensembles were selected such that members with equal amplitude values are grouped together in each subensemble. This grouping does not affect the function $f(z)$. The average of y over the subensemble should thus be identical to an average of y over the full ensemble. We assume

$$\langle y \rangle_j = \langle y \rangle, \tag{36}$$

and obtain from (35) the desired result

$$\langle xy \rangle = \langle x \rangle \langle y \rangle. \tag{37}$$

REFERENCES

1. Miller, S. E., "Coupled Wave Theory and Waveguide Applications," B.S.T.J., *33*, No. 3 (May 1954), pp. 661–719.
2. Johnson, C. C., *Field and Wave Electrodynamics*, New York: McGraw Hill, 1965.
3. Marcuse, D., "Mode Conversion Caused by Surface Imperfections of a Dielectric Slab Waveguide," B.S.T.J., *48*, No. 10 (December 1969), pp. 3187–3215.
4. Personick, S. D., "Time Dispersion in Dielectric Waveguides," B.S.T.J., *50*, No. 3 (March 1971), pp. 843–859.
5. Rowe, H. E., and Young, D. T., "Transmission Distortion in Multimode Random Waveguides," to be published in IEEE Trans. on Microwave Theory and Techniques.