the white light beam at the device output, then measuring the optical power at the four input ports. The transmission spectra from four ports, A, B, C and D, are shown in Fig. 3.

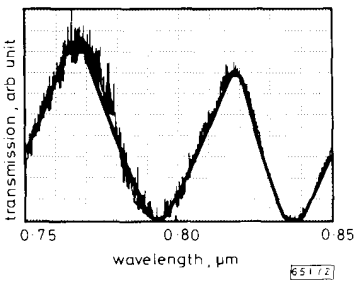

Fig. 2 *Transmission spectrum for long TMI section (4·5 mm) at 0 V*

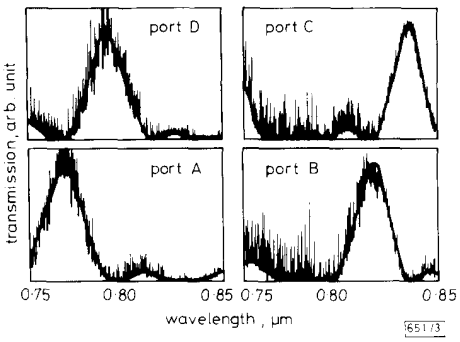

Fig. 3 *Four-channel WDM transmission spectra at ports A, B, C and D*

Each spectrum has a main transmission peak and several nulls. Ideally, the main peak will line up to the null on all other ports. The exact voltages for fine-tuning the electrodes and the exact wavelengths for four channels cannot be resolved by looking only at these spectra because of the small signal/noise ratio of this measurement. A precise measurement using laser diodes will now be explained.

The optimum wavelengths for minimising the crosstalk through the long TMI section were obtained by launching light from laser diodes into the device while thermally tuning their wavelengths. After fixing the wavelengths of the laser diodes, voltage-tuning of the short TMI sections was necessary. Because the long TMI section determines the wavelengths of the four channels, with no voltage applied to that section, we had on-hand only three out of the four laser diodes for matching four adjacent channels. The missing channel was in the range of 765 nm.

Since a diode of this kind was not available, tuning the long TMI section electrode to match to the available laser diodes was needed. Also, to minimise the voltage tuning requirements, we left one of the two short TMI sections at 0 V, and tuned the long TMI section to match it. In this case, 65 V was applied to the long TMI section. The transmission spectrum of the long TMI section is similar to Fig. 3, except that the curve is shifted towards the shorter-wavelength direction by about 20 nm. For optimum operation, 140 V was applied to the other short TMI section. The results are shown in Table 1.

Table 1 OPTICAL CROSSTALKS FOR FOUR-CHANNEL WDM

Output port	λ_1 7493 Å	λ_2 7801 Å	λ_3 8097 Å	λ_4 8339 Å
	dB	dB	dB	dB
A	−25	0	−32	−21
B	−24	−29	−30	0
C	0	−30	−18	−26
D	−21	−19	0	−33

Voltages: $V_L = 65$ V, $V_{AB} = 0$ V, $V_{CD} = 140$ V

The four channel wavelengths were found to be 749·3, 780·1, 809·7 and 833·9 nm. Optical crosstalk ranged from −18 to −33 dB.

Conclusion: A four-channel WDM comprising cascaded, dual-channel TMI WDMs was demonstrated. The operating channel wavelengths were 749·3, 780·1, 809·7 and 833·9 nm. Measured optical crosstalk ranged from −18 to −33 dB. Because laser diodes tend to mode-hop with temperature changes, the exact wavelength for perfect operation could not be easily obtained. A 1 nm deviation in wavelength from the perfect operation point may result in a degradation of 3 dB in crosstalk. In principle, all three electrodes on the device can be tuned at the same time to match the available laser diodes. Tuning the long TMI section is always preferable for locating the four channel wavelengths. Tuning of the two short TMI sections follows for obtaining the optimum crosstalk.

We thank M. Hamilton for useful discussions, W. Dougherty for photomask layout and T. Traynor for the white light source set-up.

J. P. LIN
R. HSIAO
S. THANIYAVARN

High Technology Center
Boeing Aerospace & Electronics
PO Box 3999, Seattle, WA 98124-2499, USA

7th August 1989

References

1 NEYER, A.: 'Integrated-optical multichannel wavelength multiplexer for monomode systems', *Electron. Lett.*, 1984, **20**, pp. 744–746
2 ROTTMANN, F., NEYER, A., MEVENKAMP, W., and VOGES, E.: 'Integrated-optical waveguide multiplexers on lithium niobate based on two-mode interference', *J. Lightwave Technol.*, 1988, **LT-6**, pp. 946–952
3 CHUNG, Y., YI, J. C., KIM, S. H., and CHOI, S. S.: 'Analysis of a tunable multichannel two-mode-interference wavelength division multiplexer/demultiplexer', *ibid.*, 1989, **LT-7**, pp. 766–777
4 THANIYAVARN, S., LIN, J., DOUGHERTY, W., and HSIAO, R.: 'Wavelength-polarization double multiplexed fibre optic kink', *Proc. SPIE.*, 1989, **995**, pp. 165–168

FORMANT EXTRACTION FROM PHASE USING WEIGHTED GROUP DELAY FUNCTION

Indexing terms: Signal processing, speech processing, Fourier transforms, Group delay function, Formant extraction

A method to extract formants from Fourier transform (FT) phase using a group delay (GD) function is presented. The GD function of the speech signal is difficult to process owing to the presence of large amplitude spikes. The spikes are de-emphasised by modifying the expression for computing the GD function, which facilitates formant extraction.

Introduction: Although complete information about a finite-duration signal is available in its Fourier transform (FT) phase,[1] no attempt has been made to estimate any useful parameters such as formants from the phase, because it appears to be noisy and is difficult to interpret. However, the negative derivative (with respect to the angular frequency) of the FT phase, called the group delay (GD) function, has potential applications in speech processing[2] because of its additive and high-resolution properties. The GD function of speech is, however, difficult to process owing to the presence of large amplitude spikes caused by the spectral fine structure. In this letter we present a method for formant extraction from the GD function, which involves modifying the expression for the computation of the group delay spectrum from the time signal. The effect of this modification is equivalent to weighting the group delay function by a fractional power of the FT magnitude.

Formant extraction from group delay function: A speech signal can be considered as a convolution sum of the vocal tract response and the glottal excitation signal in the time domain. The vocal tract system and the glottal excitation contribute a slowly varying component and a rapidly changing component (also called the fine structure), respectively, in the spectral domain. These components are multiplicative in the FT magnitude and additive in the group delay spectra.[2] The fine structure is due to the zeros of the transfer function introduced by the window and glottal effects, whereas the slowly varying component is due to the poles and zeros of the transfer function introduced by the vocal tract system.

The techniques used to extract formants from the FT magnitude try to capture the slowly varying component and disregard the fine structure. To derive the vocal tract characteristics from the GD spectrum, the fine structure component has to be de-emphasised. Zeros close to the unit circle appear as impulses in the GD function. These impulses form a significant part of the fine structure, and their effect cannot be completely eliminated by clipping or standard smoothing techniques.

The algorithm to compute the group delay function is given below. This is based on the expression for the derivative of FT phase as given in Reference 3.

Let $x(n)$ be the given time-domain sequence of length $N/2$ samples:

(*a*) Compute the N-point DFT $X(k)$ and $Y(k)$ of the sequences $x(n)$ and $nx(n)$ by appending the sequences with $N/2$ zeros.

(*b*) Compute the group delay function τ_p as

$$\tau_p(k) = \frac{X_R(k)Y_R(k) + X_I(k)Y_I(k)}{|X(k)|^2} \qquad k = 0, \ldots, N-1 \quad (1)$$

where the subscripts R and I denote the real and imaginary parts, respectively. If some zeros of the signal are on the unit circle in the z-plane, $\tau_p(k)$ at those frequencies might become undefined. To avoid this, we replace the denominator by a small positive value when $|X(k)| = 0$.

Fig. 1 shows a segment of speech (sampled at 10 kHz) and

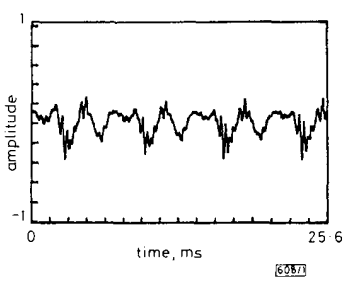

Fig. 1 *Segment of speech*

its FT magnitude, and Fig. 2 shows its FT magnitude, (linear prediction) LP and FT phase spectra. Fig. 3*a* shows the GD spectrum obtained using eqn. 1. Note that there are a number of spikes in the group delay function. The spikes are caused by the zeros close to the unit circle, which also appear as sharp valleys in the FT magnitude spectrum. The polarity and strength of a spike in the GD spectrum are determined by the location and proximity of the zero (with respect to the unit circle), respectively.[4]

To see the underlying structure of the vocal tract response from the group delay function, the effect of the spikes has to be eliminated. This is illustrated in Fig. 3*a*, where we show the lowpass-filtered GD function (thick line) where the smoothed spectrum is dominated by a large positive spike. We note that the locations of the peaks do not correspond to formants. We note from Figs. 2*a* and 3*a* that corresponding to each spike in the GD function, there is a sharp valley in the FT magnitude function, whose values are relatively very small. We observe from eqn. 1 that the square of the FT magnitude, whose values are very small at the valleys, contributes to the large spikes as it appears in the denominator. This suggests an approach to de-emphasise the spikes in the group delay by replacing the term $|X(k)|^2$ in eqn. 1 by $|X(k)|^{\alpha(k)}$, where $\alpha(k)$ varies linearly with frequency and takes on the boundary values $\alpha(0) = \alpha_1$ and $\alpha(N/2) = \alpha_2$, where $1{\cdot}0 \leq \alpha_2 < \alpha_1 < 2{\cdot}0$.

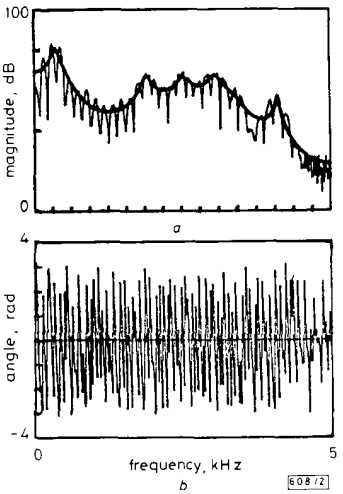

Fig. 2 *Magnitude and phase spectra of speech*

a Fourier transform magnitude spectrum and its smoothed envelope (linear prediction spectrum)
b Fourier transform phase spectrum

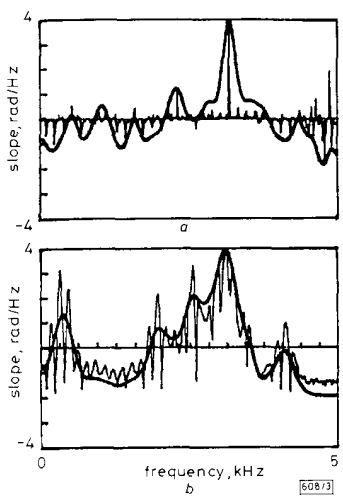

Fig. 3 *Group delay spectrum derived from speech segment*

a Original group delay spectrum and its smoothed envelope
b Modified group delay and its smoothed envelope

This term is made a function of frequency to adjust for the large dynamic range in $\tau_p(k)$ due to formants.

Fig. 3*b* shows the modified group delay spectrum. In our present study we found that the values $\alpha_1 = 1{\cdot}25$ and $\alpha_2 = 1{\cdot}0$ seem to give good results for a large number of segments. The values of α_1 and α_2 are not very critical as long as they are close to 1. We observe that the strength of the spikes has been significantly reduced relative to the neighbouring values. The thick line in Fig. 3*b* shows the smoothed group delay spectrum. On comparison with the thick line in Fig. 2 we note that the location of peaks matches with that of the LP spectrum peaks.

Conclusions: In this letter we describe a procedure to extract formants from FT phase by processing a modified group delay function. The peaks and valleys in the smoothed modified group delay correspond to formants and natiformants, respectively. Our experiments with synthetic and natural speech

data have demonstrated that formants can be estimated consistently and reliably from the smoothed modified group delay function.

H. A. MURTHY
K. V. MADHU MURTHY*
B. YEGNANARAYANA

Department of Computer Science & Engineering
Indian Institute of Technology
Madras 600 036, India

* *Department of Computer Applications*
Sri Venkateswara University College of Engineering
Tirupati 517 502, India

31st July 1989

References

1 OPPENHEIM, A. V., and LIM, J. S.: 'The importance of phase in signals', *Proc. IEEE*, 1981, **69**, pp. 529–541
2 YEGNANARAYANA, B.: 'Formant extraction from linear prediction phase spectrum', *J. Acoust. Soc. Am.*, 1978, **63**, pp. 1638–1640
3 OPPENHEIM, A. V., and SCHAFER, R. W.: 'Digital signal processing' (Prentice-Hall, Englewood Cliffs, NJ, 1975), pp. 495–498
4 YEGNANARAYANA, B., SAIKIA, D. K., and KRISHNAN, T. R.: 'Significance of group delay functions in signal reconstruction from spectral magnitude or phase', *IEEE Trans.*, 1984, **ASSP-32**, pp. 610–622

ANALYSIS OF FORWARD WIDE-ANGLE LIGHT PROPAGATION IN SEMICONDUCTOR RIB WAVEGUIDES AND INTEGRATED-OPTIC STRUCTURES

Indexing terms: Waveguides, Optical waveguides, Optoelectronics, Integrated optics

We analyse longitudinally varying semiconductor rib waveguides with a new nonparaxial wide-angle equation for unidirectional light propagation. We further develop a solution method involving multiplication of the incoming electric field by a series of unitary operators which are evaluated by split-step fast Fourier transform and finite difference techniques. In the test case of a strongly guiding rib waveguide Y-junction, the estimated losses nearly coincide with those of standard Fresnel equation procedures. We also present a numerical analysis of an integrated-optic lens, demonstrating significant errors in previous results. Finally, we explicitly illustrate the increased accuracy of our new method in comparison to the Fresnel equation for a highly nonparaxial Gaussian beam.

Introduction: Earlier numerical studies of the differential losses of semiconductor rib waveguide Y-junctions[1–3] have been restricted to the Fresnel equation. While the paraxial approximation was proven in these studies to agree in the weakly guiding limit with the standard effective index method[4] and in the strongly guiding limit with experimental measurements,[5] no theoretical treatment of the errors associated with nonparaxial radiation modes has yet been advanced. Accordingly, we here derive a new wide-angle wave equation applicable to general waveguiding structures. We then introduce a solution procedure incorporating both fast Fourier transform and finite difference methods, and apply this method to the computation of the differential loss of a strongly guiding and longitudinally varying semiconductor rib waveguide. We also repeat this calculation with a paraxial algorithm that does not require matrix inversions or Fourier transformations.[6] The results are in excellent agreement with those of the wide-angle formulation. We further consider a fisheye lens, which we find can similarly be analysed with the Fresnel equation. Finally, we consider the propagation of a highly nonparaxial Gaussian beam in a homogeneous medium. This example clearly illustrates the greater range of applicability of the wide-angle equation in comparison to the Fresnel equation.

Numerical method: Derivations of wide-angle wave equations generally proceed from the following expression for a single polarisation component of a forward-propagating monochromatic electric field:

$$E(x, y, z + \Delta z) = e^{\delta\sqrt{(1 + X + Y)}}E(x, y, z) \tag{1}$$

We have introduced $\delta = -ik_0 n_0 \Delta z$ and $X = X_0 + N$ with $N = [(n^2/n_0^2) - 1]$, $X_0 = (1/k_0 n_0)^2(\partial^2/\partial x^2)$ and $Y = (1/k_0 n_0)^2(\partial^2/\partial y^2)$. The variables k_0, n_0 and Δz denote the vacuum wavevector, a typical guided mode refractive index and the longitudinal step length, respectively. Unlike previous studies in which the square-root operator in eqn. 1 was written in polar co-ordinates and only the radial term expanded to second-order,[7] we here write

$$e^{\delta[\sqrt{(1 + X + Y)} - 1]} \simeq e^{(\delta/2)[X - (X^2/4)]} \times e^{-(\delta/8)(NY + YN)}e^{(\delta/2)[Y - (Y^2/4) - (X_0 Y/2)]} \tag{2}$$

which may be approximated still to second-order in X and Y by[7]

$$\left(\frac{1 + \left(\frac{1 + \delta}{4}\right)X}{1 + \left(\frac{1 - \delta}{4}\right)X}\right)\left(\frac{1 + \frac{\delta}{16}(NY + YN)}{1 - \frac{\delta}{16}(NY + YN)}\right) \times e^{(\delta/2)[Y - (Y^2/4) + (X_0 Y/2)]} \tag{3}$$

The first two and the last of these operators are then evaluated using a standard finite difference procedure and the split-step fast Fourier transform algorithm, respectively. Note that if we instead define $X = X_0 + N/2$ and $Y = Y_0 + N/2$, we obtain a fully symmetric expression in place of eqn. 2 which can be analysed in a similar fashion.

An alternative Fresnel equation approach which has not previously been applied to optical propagation problems involves direct multiplication by unitary band-diagonal matrices,[6] and may be formulated as follows:

$$E(x, y, z + \Delta z) = Y^{(2)}Y^{(1)}Y^{(2)}X^{(2)}X^{(1)}X^{(2)} \times \exp\left[-\frac{ik_0}{2n_0}\int_z^{z+\Delta z}(n^2(x, y, z') - n_0^2)\,dz' + 4i(A_X^{(1)} + A_Y^{(2)})\right] \times Y^{(2)}Y^{(1)}Y^{(2)}X^{(2)}X^{(1)}X^{(2)}e^{-ik_0 n_0 \Delta z}E(x, y, z) + O(\Delta z)^3 \tag{4}$$

Assuming an even number of grid points N_w separated by a distance Δw in the w-direction, where $w \in \{x, y\}$, $W^{(L)}$ with $W \in \{X, Y\}$ and $L \in \{1, 2\}$ is the symmetric and unitary block diagonal matrix with nonzero upper elements given by $W^{(L)}_{j,j} = \cos A_W^{(L)}$ for $j = 3 - L, 4 - L, \ldots, N_w + L - 2$ and $W^{(L)}_{2j+1-L,2j+2-L} = -i \sin A_W^{(L)}$ for $j = 1, 2, \ldots, (N/2) + L - 2$ and $A_W^{(L)} = \Delta z/(4k_0 n_0 L(\Delta w)^2)$. While each propagation step is now very rapid, small step lengths are essential.

Results: We have first considered the effect of nonparaxiality on light propagation in integrated-optic rib waveguides. In particular, we considered a rib waveguide 1·1 μm high and 2 μm wide etched into a 1·3 μm-thick epilayer with refractive index $n_1 = 3{\cdot}44$. The cladding refractive index is $n_2 = 3{\cdot}34$ for which the normalised modal refractive index at our light wavelength $\lambda = 1{\cdot}55\,\mu$m equals $b \equiv (N^2 - n_2^2)/(n_1^2 - n_2^2) \simeq 0{\cdot}498$. The refractive index profile of our 40 μm-long Y-junction is given by the union of two such waveguide profiles centred at $\pm x(z)$, where $x(z) = 1 - \cos(\pi z/40\,\mu\text{m})$. Our analysis employs a computational window with $L_x \times L_y = 8 \times 4\,\mu\text{m}^2$ and $N_x \times N_y = 64 \times 64$ transverse grid points. Further, to absorb the radiated light we add $-5{\cdot}63i$ and $-2{\cdot}81i$ to the squared refractive index of the grid points at the computational window edge and of the immediately adjacent points, respectively.[1,2,8]