Corning: Inverse Problems for Glass

Greg Luther

Eyal Arian, Andrew Lacey, John McCuan, Peter Monk, John Pelesko, Rakesh, Colin Please, Chris Raymond, Don Schwendeman, Tom Witelski

Abstract

Three individual glass-related topics were discussed during the workshop. The first, how to obtain the refractive index at points inside a transparent body, received most attention. This was a two-dimensional problem, as the objects of interest were glass fibres, or their pre-forms, and involved the use of measurements of positions and directions of outgoing beams of light when given beams are shone in. A number of numerical and analytical approaches were discussed.

The second was a one-dimensional grating-design problem: how to get desired transmitted and reflected light intensity profiles or distributions by choosing an appropriate distribution of refractive indices along a fibre. An optimization approach was developed: the direct problem was solved for different frequencies of light and a measure of the difference between the resulting light outputs and those desired was minimized.

The third and last was again experimentally related and again one-dimensional. Here, the use of measured light, for instance its polarization, being re-emitted from the end of a fibre through which a pulse of light had originally been input, towards finding fibre properties, was investigated. There was some consideration of the model for light reflection; no improvements on the methods suggested in the literature for determining coefficients in the model were forthcoming.

1 Determination of Refractive Index

For this inverse problem of determining the refractive index $n(x, y)$ inside a transparent cylindrical body, specifically a glass fibre or its pre-form, two main techniques were considered. The first was based on looking at a reduction of the Helmholtz equation representing the electric field of light travelling through the object in the case of small variation of refractive index, (a situation of genuine interest): §1.3. Although the problem was posed, no real progress was made towards its solution, even for special geometries such as circular symmetry.

More work was done, and progress towards methods of solution achieved, using geometric optics (valid if the wavelength of light used is small compared with geometrical length scales). Specifically, a numerical method was developed for cases where the refractive index varies little; §1.4.1. (Another technique is suggested in §1.4.3.) Problems with distinct media, so n is essentially stepwise constant with just thin transition zones, were investigated more thorougly (§1.4.2), only polygonal geometries with small jumps of refractive index were discussed in detail. A numerical method was developed for this last special case. Finally, an analytic solution for radially symmetric problems was obtained, even for significant variation of refractive index; §1.4.3.

1.1 Introduction

This problem concerned the determination of the refractive index inside an optical fibre (with a width of order 1 mm) or its "pre-form" (with a width of order 1 cm). The general method is to shine a beam of light through the effectively two-dimensional object under investigation and measure the intensity, location and/or direction of the transmitted beam. Using these measurements with beams incident from a variety of directions, the refractive index $n(x, y)$ is to be determined.

The equations satisfied by the electric field $\mathbf{E}(x, y)e^{-i\omega t}$ (taking just two dimensions) of a beam of light in the glass are:

$$\frac{\partial}{\partial t}\left(\epsilon\frac{\partial}{\partial t}(\mathbf{E}e^{-i\omega t})\right) = -\omega^2\epsilon\mathbf{E}e^{-i\omega t} = \frac{1}{\mu_0}\nabla\times\mathbf{B} = -\frac{1}{\mu_0}\nabla\times\left(\nabla\times(\mathbf{E}e^{-i\omega t})\right)$$
$$= \frac{1}{\mu_0}(\nabla^2\mathbf{E} - \nabla(\nabla\cdot\mathbf{E}))e^{-i\omega t}$$

with $\nabla\cdot(\epsilon\mathbf{E}e^{-i\omega t}) = (\epsilon\nabla\cdot\mathbf{E} + \mathbf{E}\cdot\nabla\epsilon)e^{-i\omega t} = \mathbf{0}$. Then

$$\nabla^2\mathbf{E} + \epsilon\mu_0\omega^2\mathbf{E} + \nabla\left(\frac{\nabla\epsilon}{\epsilon}\cdot\mathbf{E}\right) = \mathbf{0}\,,$$

where $\epsilon = \epsilon_r(x, y)\epsilon_0$, ϵ_r = relative permittivity, $\epsilon_0\mu_0 = 1/c^2$, c = speed of light in a vacuum, $\omega/c = k = 2\pi/\lambda$, k = wave number, λ = wavelength in vacuum and $\epsilon_r = 1/n(x, y)^2$.

For light polarised in the z-direction, this reduces to

$$\nabla^2 E + k^2 n^2 E = 0. \tag{1}$$

(The basic ray theory for the other polarisation is similar to that for this simpler case.)

The inverse problem to be solved is that of determining $n(x, y)$, given an incident electrical field, specified as that of a plane wave or a narrow beam (or ray) of light, and measured transmitted electromagnetic radiation (amplitude and/or direction, not the phase). Two limiting cases were of possible interest:

(i) Small variation of refractive index. The present technique entails immersing the body in a liquid with refractive index close to that of the glass(es). Then n is likely to vary at most by 10% and possibly as little as 0.2%. (Ideally the method could be employed in air, meaning that n would change more substantially at the outer boundary.)

(ii) High wave number. For k large, in comparison with the dimensions of the body (and with length scales of internal variation), geometrical optics can be used. For a pre-form, the internal length scale is typically millimetres compared with a wavelength of about 1 μm, giving a nondimensional k of about 3000.

During the workshop, the second was taken to be of primary importance, although the smallness of the variation of n was briefly looked at in isolation and considerable attention was paid to cases of n changing little at the same time as having k large.

1.2 Earlier work

In the 1995 workshop, the special case of radial symmetry, with large wave number and small variation of refractive index, had been dealt with. The starting point for this, and the geometrical optics theory of §1.4, is to try the ansatz

$$E = A(\mathbf{x})e^{iku(\mathbf{x})} \tag{2}$$

as a solution of (1). Regarding k as large,

$$(-k^2|\nabla u|^2 A + 2ik\nabla u\cdot\nabla A + ikA\nabla^2 u)e^{iku} + k^2n^2Ae^{iku} = 0 \tag{3}$$

so, to leading order, u (which gives the phase of the electromagnetic wave) satisfies the Eikonal equation

$$|\nabla u|^2 = n(x,y)^2\,. \tag{4}$$

This can be solved by Charpit's equations, writing (4) as

$$\frac{1}{2}|\nabla u|^2 - \frac{1}{2}n^2 = F(x,y,p,q), \qquad p = \frac{\partial u}{\partial x}, \qquad q = \frac{\partial u}{\partial s},$$

leading to

$$\frac{dx}{d\hat{s}} = \frac{\partial F}{\partial p} = p\,, \quad \frac{dy}{d\hat{s}} = \frac{\partial F}{\partial q} = q, \quad \frac{dp}{d\hat{s}} = -\frac{\partial F}{\partial x} - p\frac{\partial F}{\partial u} = n\frac{\partial n}{\partial x}, \quad \frac{dq}{d\hat{s}} = -\frac{\partial F}{\partial y} - q\frac{\partial F}{\partial u} = n\frac{\partial n}{\partial y},$$

$$\text{and} \qquad \frac{du}{d\hat{s}} = p\frac{\partial F}{\partial q} + q\frac{\partial F}{\partial q} = p^2 + q^2,$$

with $\hat{s}$ being a parameter along rays. (The rays are of course perpendicular to the surfaces of constant phase, $u =$ const.) Reparameterising the rays, using arc length s,

$$\frac{d\hat{s}}{ds} = \left(\left(\frac{dx}{d\hat{s}}\right)^2 + \left(\frac{dy}{d\hat{s}}\right)^2\right)^{-\frac{1}{2}} = (p^2+q^2)^{-\frac{1}{2}} = \frac{1}{n}\,,$$

so

$$\frac{dx}{ds} = \frac{p}{n}\,, \quad \frac{dy}{ds} = \frac{q}{n}\,, \quad (p^2+q^2 = n^2)\,, \quad \frac{dp}{ds} = \frac{\partial n}{\partial x}\,, \quad \frac{dq}{\partial s} = \frac{\partial n}{\partial y}\,, \quad \text{and} \quad \frac{dy}{ds} = n\,. \tag{5}$$

The first of these, on eliminating p and q, gives equation (4) of the 1995 report:

$$\frac{d}{ds}\left(n\frac{d\mathbf{x}}{ds}\right) = \nabla n\,. \tag{6}$$

If polar coordinates (r,θ) are used instead, $\dfrac{dx}{dr} = \left(\dfrac{dr}{ds}, r\dfrac{d\theta}{ds}\right)$ and (6) becomes

$$\left(n\frac{d^2r}{ds^2} + \frac{\partial n}{\partial r}\left(\frac{dr}{ds}\right)^2 + \frac{\partial n}{\partial \theta}\frac{dr}{ds}\frac{d\theta}{ds} - nr\left(\frac{d\theta}{ds}\right)^2, \frac{\partial n}{\partial r}r\frac{dr}{ds}\frac{d\theta}{ds} + \frac{\partial n}{\partial \theta}r\left(\frac{d\theta}{ds}\right)^2 + 2n\frac{dr}{ds}\frac{d\theta}{ds} + rn\frac{d^2\theta}{ds^2}\right)$$

$$= \left(\frac{\partial n}{\partial r}, \frac{\partial n}{\partial \theta}\right)\,.$$

Should the body be radially symmetric, $n = n(r)$, this simplifies to

$$\left(n\frac{d^2r}{ds^2} + \frac{dn}{dr}\left(\frac{dr}{ds}\right)^2 - nr\left(\frac{d\theta}{ds}\right)^2, \frac{1}{r}\frac{d}{ds}\left(nr^2\frac{d\theta}{ds}\right)\right) = \left(\frac{dn}{dr}, 0\right).$$

It follows that

$$nr^2\frac{d\theta}{ds} = \text{const.} = -n_0 Y \tag{7}$$

and, since $\left(\frac{dr}{ds}\right)^2 + \left(r\frac{d\theta}{ds}\right)^2 = 1$,

$$\left(\frac{dr}{ds}\right)^2 = 1 - \frac{n_0^2 y^2}{n^2 r^2}. \tag{8}$$

These are equations (7) and (6) of the 1995 report.

Returning to (6), if the ray is initially parallel to the x-axis, say $y = Y$ (or even nearly so) and n only has small variation, to leading order $y \equiv Y$. Because $x \sim s$, the "paraxial approximation", given by the y component of (6), determines the first-order correction:

$$n_0\frac{d^2y}{dx^2} \sim \frac{\partial n}{\partial y}, \tag{9}$$

where n_0 = refractive index away from the pre-form (and $n \sim n_0$ everywhere). In the special case of radial symmetry this is just equation (8) of 1995:

$$\frac{d^2y}{dx^2} = \frac{y}{rn_0}\frac{dn}{dr}. \tag{10}$$

Integrating (10), once, along the ray (*i.e.* along $y = Y$, approximately), say from $-X$ to X, X taken large enough so that the pre-form lies inside $r < X$ (X can be assumed small enough to be closer to the pre-form than the light source and detector are), gives the change in direction:

$$\begin{aligned}\phi \sim \tan\phi &\sim \int_{-X}^{X} \frac{Y\frac{dn}{dr}(r)}{rn_0}dx && (r^2 = x^2 + Y^2)\\ &= \frac{2Y}{n_0}\int_Y^\infty \frac{\frac{dn}{dr}dr}{\sqrt{r^2 - Y^2}} && (n \equiv n_0 \text{ for large enough } r).\end{aligned}$$

This "Abel integral equation", can be written as

$$\phi(Y) = \frac{2Y}{n_0}\int_Y^\infty \frac{\frac{d}{dr}(n(r) - n_0)}{(r^2 - Y^2)^{\frac{1}{2}}}dr \tag{11}$$

and has solution

$$n(r) - n_0 = -\frac{n_0}{\pi}\int_r^\infty \frac{\phi(Y)dY}{(Y^2 - r^2)^{\frac{1}{2}}} \tag{12}$$

(see eqs (11) - (14) of 1995). (The equation (11), or the solution (12), can be converted to a more standard form as in [5], [9], [14], [20] by a straightforward change of variables; although the manipulation is not quite trivial. The solution to the standard form is derived via fractional integration in [5] and [14]. It is possible to check, in substituting (11) into (12) and vice versa, and noting that $n(r) - n_0 = 0$ at $r = \infty$ and $\phi(Y) = 0$ at $Y = \infty$, that (12) gives a, and the only, solution to (11).)

The suggested technique is then to measure the angle of deflection $\phi(Y)$ of a narrow beam passing a distance Y from the centre (of the axis) of the object. The inversion (12) gives the refractive index.

1.3 Nearly constant refractive index

Taking plane incident light (a plane polarised electromagnetic field), say e^{in_0kx} for light of unit amplitude propagating in the x direction, (1) can be used, at least in principle, to determine the outgoing wave (transmitted and reflected light) in terms of $n(x, y)$. It is, of course, the inverse problem, finding n from knowledge about the transmitted light, which we really want to solve.

The electromagnetic field can be written as the sum of the incident wave and an outgoing wave resulting from the presence of the pre-form:

$$E = e^{in_0kx} + v(x, y).$$

Here v satisfies

$$\nabla^2 v + k^2 n^2 v + k^2(n^2 - n_0^2)e^{in_0kx} = 0 \tag{13}$$

with the Sommefeld radiation condition

$$\frac{\partial v}{\partial r} - in_0kv = O(v) \qquad \text{for} \qquad r \to \infty. \tag{14}$$

For small variation of n, it is possible to write $n = n_0 + \delta N(x, y)(N \equiv 0$ outside the pre-form) with $0 < \delta \ll 1, v = \delta V$, and, to leading order,

$$\nabla^2 V + k^2 n_0^2 V + 2k^2 n_0 N e^{in_0kx} = 0.$$

This (direct) problem can be solved, subject to (14), in terms of the Hankel function

$$H_0^{(1)}(z) = J_0(z) + iY_0(z)$$

$$\left(H_0^{(1)}(z) \sim \sqrt{\frac{2}{\pi z}} e^{i(z-\pi/4)} \text{ for } z \to +\infty \,, \quad H^{(1)} \sim \frac{2i}{\pi} \ln z \quad \text{ for } z \to 0+ \right):$$

$$\begin{aligned} V(\mathbf{x}) &= \tfrac{1}{4} \textstyle\int\int H_0^{(1)}(kn_0|\mathbf{x} - \boldsymbol{\xi}|) 2k^2 n_0 N(\xi, \eta) e^{in_0k\xi} \, d\xi \, d\eta \\ &= \frac{ik^2 n_0}{2} \int\int H_0^{(1)}(kn_0|\mathbf{x} - \boldsymbol{\xi}|) N(\boldsymbol{\xi}) e^{in_0k\xi} \, d\xi \, d\eta \,. \end{aligned} \tag{15}$$

This, then, is the approximate solution to the direct problem. The inverse problem, the one of real interest, amounts to solving the integral equation (15) with $V(x, y)$ specified (by measurements in experiments). (V would not be fully specified; for instance, measurements would not give the phase of V.) This was not done during the workshop.

One possible way of making progress might be thought to now use the large size of the wave number k. From the asymptotic behaviour of the Hankel function for large argument, (15) becomes

$$V(\mathbf{x}) \sim \frac{ik^{3/2} n_0^{1/2}}{\sqrt{2\pi}} \int\int \frac{e^{i(nk_0(|\mathbf{x}-\boldsymbol{\xi}|))-\frac{\pi}{4})}}{\sqrt{|\mathbf{x}-\boldsymbol{\xi}|}} N(\boldsymbol{\xi})\, d\xi\, d\eta\,. \tag{16}$$

This does not immediately appear at all similar to the results obtained below using the primarily ray theory/geometric optics (largeness of k) and secondly the weak variation of refractive index (smallness of δ). One reason is the way the field is written: here, E = original wave + the perturbation δV; below, the field is a single wave. Another, and related, reason is a non-uniformity in the small-δ approximation when k is large: for instance, $e^{inkx} \sim e^{in_0kx}(1 + i\delta Nkx + \dots)$ is an invalid expansion for k of size $1/\delta$. (Taking, for simplicity, $y = 0$, and looking at "large" distances, $x \gg 1$ – so that a genuine outgoing wave can be expected – $|\mathbf{x}-\boldsymbol{\xi}|+\xi \sim x+\eta^2/2x+\dots$. Substituting this into the exponential in (16), and assuming that $N \equiv$ constant in the pre-form, leads to

$$V \sim C_1 k^{\frac{3}{2}} x^{-\frac{1}{2}} e^{in_0kx} \int\int e^{in_0k\eta^2/x}\, d\xi\, d\eta\,.$$

Carrying out the η integration appears – without checking details – to make

$$V \sim C_2 k e^{in_0kx}\,.$$

This is an expected transmitted wave, or rather its perturbation, and is seen to lead to the nonuniformity if k is of order $1/\delta$.)

1.4 Geometric optics

If k is large, in the sense that the length scale of the object under investigation is much larger than the wavelength, the ansatz (2) makes sense and Eikonal equation (4) for the phase can be derived. Ray theory, essentially solving (6), can then be used in determining deflections of beams of light.

Three overlapping special cases were looked at during the workshop.

(1) Small variation in refractive index. Looking primarily at high wave number and then at n changing only a little appears to be easier than the reverse (just above).

(2) Distinct media. Here n is piecewise constant. Specifically, the case of n taking only two values, say an outer value n_0 and n_1 in an inner material was considered. This is a case of some practical interest, as a pre-form can be made from two types of glass with an only very thin zone where one blends into the other.

(3) Radial symmetry, so $n = n(r), r = \sqrt{x^2 + y^2}$.

The case dealt with in 1995, §1.2 here, is a combination of *(1)* and *(3)*. For *(2)*, Snell's law applies at the interfaces where n jumps:

$$n_I \sin\varphi_I = n_T \sin\varphi_T$$

for φ_I = angle of incidence, φ_T = angle of transmission, n_I = refractive index of first medium and n_T= that of the second; see Fig.1.

To date, the reflected rays have not been considered. With *(1)*, reflected waves have amplitude $O(\delta)$, if δ is a measure of the small variation of refractive index. The power associated with reflected waves is then $O(\delta^2)$, whereas in this report we only work to $O(\delta)$, at most.

(Of course Snell's law with the reflected wave, can be derived from the original Maxwell's equations, using appropriate continuity conditions at the surface. It can also be obtained by integrating, say the 3rd of (5) across a narrow zone, across which n changes from n_I to n_T, and in which $n \sim n(ax+by)$ so that the interface is locally $x\cos\alpha + y\sin\alpha =$ const., $a = c\cos\alpha, c\sin\alpha$. It helps to write $p = n\cos\theta, q = n\sin\theta$, θ giving the direction of the ray.)

Also at present, total internal reflection has not been allowed for, and the failure of geometric optics to deal with rays glancing a surface where refractive index is discontinuous has not been considered. When the refractive index varies little, $n = n_0 + \delta N(x,y)$ with $\delta \ll 1$, to have total internal reflection, rays must be incident on surfaces at angles to the normal of nearly $\pi/2$ and the awkward cases both exceptional, and roughly identical.

As in 1995, in what follows, we shall regard the output of the experiment as a measurement of the position and/or direction of a departing beam, resulting from a known, narrow source: this knowledge of a family of rays is to determine the refractive index. There is an equivalence between this sort of measurement and that of intensity of transmitted light from a broader beam, for example if the incoming light is a plane wave. If we return to the ansatz (2) but now consider the $O(k)$ terms in (3), the amplitude of the electric field is seen to satisfy, approximately,

$$\nabla u \cdot \nabla A = -\frac{1}{2} A \nabla^2 u\,.$$

Along characteristics of this equation (the rays),

$$\frac{dA}{d\hat{s}} = -\frac{1}{2}\nabla^2 A$$

so, for an initial amplitude of unity,

$$A = \exp\left(-\frac{1}{2}\int \nabla^2 u\, d\hat{s}\right) = \exp\left(-\frac{1}{2}\int \frac{1}{h_1 h_2 h_3}\frac{\partial}{\partial X_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial u}{\partial X_1}\right) d\hat{s}\right),$$

where we have changed to orthogonal co-ordinates (X_1, X_2, X_3) with $X_1 = u$ and $h_j \equiv |\partial \mathbf{x}/\partial X_j|$.

From

$$\left|\frac{\partial \mathbf{x}}{\partial X_1}\right| = \frac{1}{du/ds} = \frac{1}{n},$$

(5), $\partial/\partial X_1 = n\partial/\partial s$, $d\hat{s} = ds/n$, and the identification of $h_2 h_3$ with the area of a small element, say $\mathcal{S}$,

$$\frac{1}{h_1 h_2 h_3}\frac{\partial}{\partial X_1}\left(\frac{h_2 h_3}{h_1}\frac{\partial u}{\partial X_1}\right) = \frac{1}{\mathcal{S}}\frac{d}{ds}(n\mathcal{S})$$

and

$$A = \exp\left(-\frac{1}{2}\int \frac{d}{ds}(n\mathcal{S})ds/n\mathcal{S}\right) = \sqrt{\frac{n_0 \mathcal{S}_I}{n\mathcal{S}}}.$$

The final power, which is proportional to A^2, is then $\mathcal{S}_I/\mathcal{S}_O$, on taking the initial power to be unity, where $\mathcal{S}_O$ is the output area bounded by the rays which also bound the input area $\mathcal{S}_I$. For

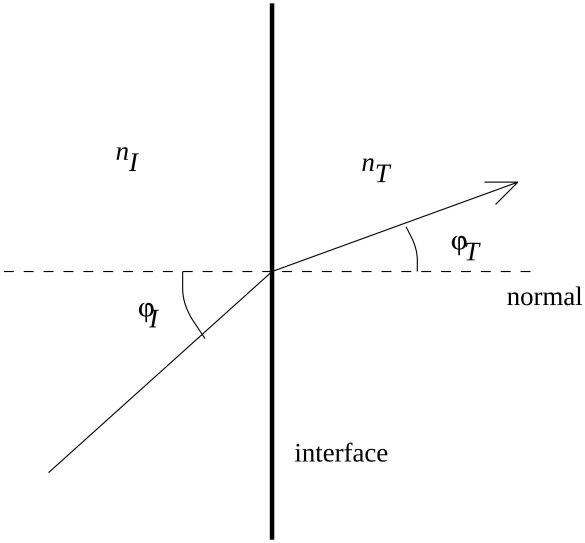

Figure 1: Angles in Snell's law.

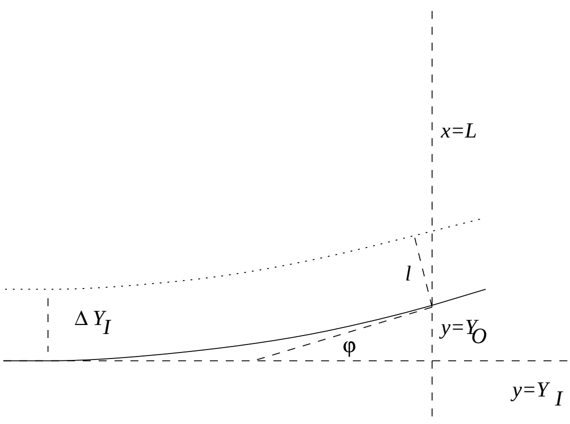

Figure 2: Deviation in a ray with its changes of distance (equivalent to areas bounded by rays).

the two-dimensional problem under discussion here, (relative) final intensity $= l/\Delta Y_I$, with l the distance shifted, perpendicular to the final rays, when the initial ray is translated through ΔY_I.

For the final rays making an angle φ with the x direction, see Fig.2, the power per unit length at the measuring device is then

$$\frac{l}{\Delta Y_I}\cos\theta = \frac{\Delta Y_O}{\Delta Y_I} = \frac{1}{dY_O/dY_I}\,.$$

1.4.1 Small variation of n.

Here, as in the special case discussed in 1995, the ray equation (6) can be reduced to (9) – taking axes orientated so that the original ray is in the x direction. If the light source is assumed to be at $x = -L$, $y = Y_I$ and the detector is at $x = L$, one integration of (9) yields the direction,

$$\frac{dy}{dx} \sim \frac{1}{n_0}\int_{-L}^{x} \frac{\partial n}{\partial y}(\xi, y)\, d\xi\,,$$

and a second integration gives the final position of the beam:

$$Y_O \sim Y_I + \frac{1}{n_o}\int_{-L}^{L} (L-x)\frac{\partial n}{\partial y}(x, Y_I)\, dx\,.$$

Rather than work with $\partial n/\partial y$, a further integration can be performed, this time with respect to y, to get an equation of the form

$$\int_{-L}^{Y_I} (Y_O(y) - y)\, dy = \frac{1}{n_0}\int_{-L}^{L} (L-x)(n(x,y) - n_0)\, dy\,. \tag{17}$$

This equation, with suitable modifications for changes of axes, also holds for other ray directions. The problem can be discretized, with the integrals being replaced by summations, and a system of linear equations for $n_{ij} = n(x_i, y_j)$ is obtained. Fig.3 shows solutions obtained for an elliptical object. In this test case, n was taken to be a constant, n_1, different from n_0, in $(x/0.3)^2 + (y/0.6)^2 < 1$. A set of data for $Y_O(y)$ was generated for uniformly distributed, parallel incident rays, all for a succession of 40 equally spaced angles between 0 and 2π. Peter Monk's code for the numerical solution of (17) was then used to determine n from this data. Note the better results for the coarser discretization.

An alternative approach uses the earlier symmetric analysis, see §1.4.3.

1.4.2 Distinct media.

A particularly simple situation is when there is a region Ω occupied by glass of refractive index $n_1 \neq n_0$; outside Ω the refractive index $\equiv n_0$. If a family of parallel beams are employed, the extreme undeflected ones, see Fig.4(a), bound the region Ω.

Carrying out this procedure for a variety of directions will identify the convex hull of Ω. If Ω is known to be convex, the envelope of these rays will actually be the surface $\partial\Omega$ of the distinct material, see Fig.4(b). This procedure should work whether or not n_1 is known, or constant. In principle, once the outer boundary (or the covex hull) of Ω has been found by such a technique, if n_1 is known (and the boundary is smooth) the trajectory of rays as they enter (and just as they

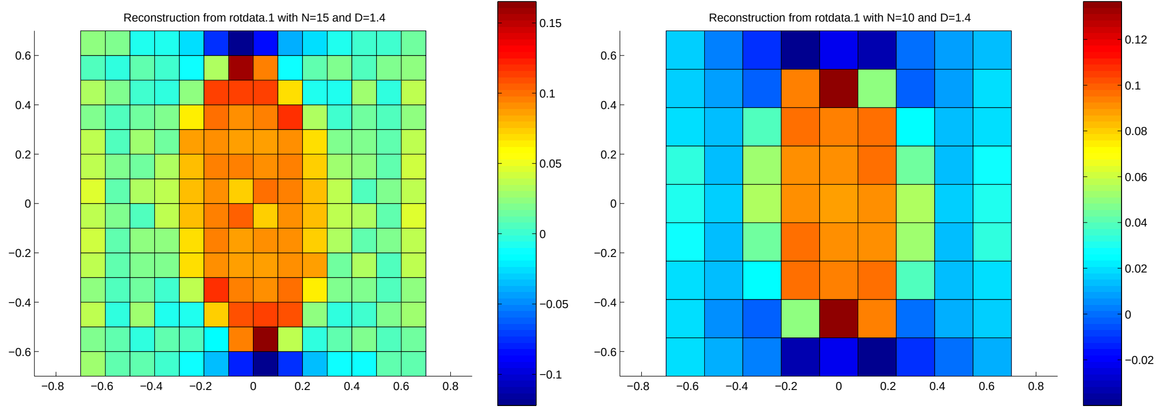

Figure 3: The refractive index, according to colour, obtained by solving discrete versions of (17). The left-hand plot uses a 14×14 discretization while the right-hand one is 9×9; both employ the same data.

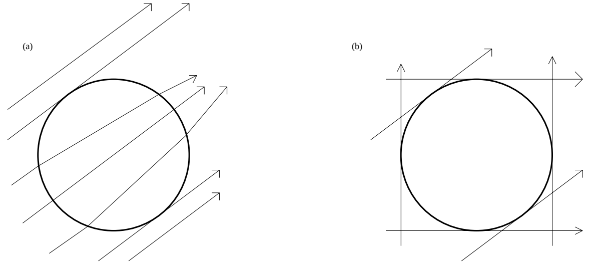

Figure 4: The "shadow method".

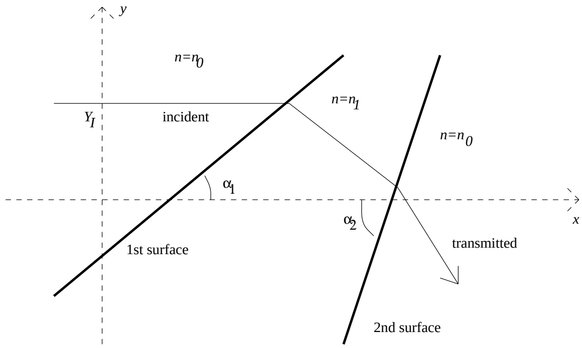

Figure 5: A light ray crossing a polygonal, or slab, inclusion.

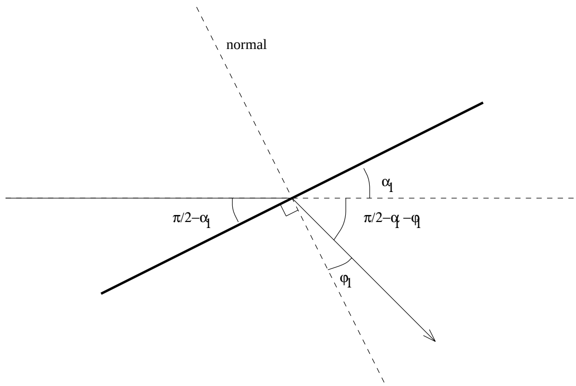

Figure 6: Change of direction of the ray at the first surface, $y = x \tan \alpha_1 + \beta_1$.

leave) Ω can be found. This would allow the the method to be used again, or the technique to be combined with the method below – if the interior variation is small – to identify a secondary inclusion.

Another special case is that of a polygonal inclusion, with known constant refractive index $n_1 \neq n_0$. Assuming that rays pass through the boundary twice (which will be true for a convex inclusion), the problem becomes one of identifying two straight lines (really surfaces), for the present represented by $y = x \tan \alpha_1 + \beta_1$, $y = x \tan \alpha_2 + \beta_2$, from the final paths of the light, see Fig.5.

Using Snell's law, if the incident beam is parallel to the x axis, more specifically the ray is $y = Y_I$, the intermediate beam (where $n = n_1$) makes an angle $\pi/2 - \alpha_1 - \varphi_1$, where $\varphi = \sin^{-1}(\frac{n_0}{n_1} \cos \alpha_1)$, with the x direction – measured clockwise; see Fig.6.

The intermediate ray is then

$$y - y_1 = -(x - x_1) \cot(\varphi_1 + \alpha_1) \, ,$$

where $y_1 = Y_I$ and $x_1 = (y_1 - \beta_1) \cot \alpha_1$.

Another application of Snell's law, at (x_2, y_2), where the beam, or ray, crosses the second surface

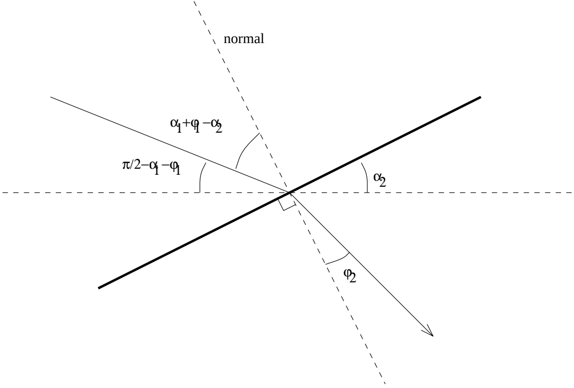

Figure 7: Change of direction of the ray at the second surface, $y = x\tan\alpha_2 + \beta_2$.

back into the region of $n = n_0$, leads to the final, transmitted, ray

$$(y - y_2) = -(x - x_2)\cot(\alpha_2 + \varphi_2)\,,$$

see Fig.7, where

$$y_2 = y_1 - (x_2 - x_1)\cot(\alpha_1 + \varphi_1) = x_2\tan\alpha_2 + \beta_2 \qquad \text{and } \varphi_2 = \sin^{-1}\left(\frac{n_1}{n_0}\sin(\alpha_1 + \varphi_1 - \alpha_2)\right)\,.$$

A measurement of the location and direction of the transmitted ray gives two pieces of information for the four quantities α_j, β_j so other rays must also be used. A parallel beam ($y \equiv \hat{Y}_I \neq Y_I$) would produce a final ray with the same direction as previously, leading to only one more item of information, so another angle must be employed. The use of a beam with direction θ, $y = \hat{Y}_I + x\tan\theta$, is equivalent to rotating the fibre through θ in the clockwise direction – or using new co-ordinates $\tilde{x}$, $\tilde{y}$,

$$x = \tilde{x}\cos\theta - \tilde{y}\sin\theta\,, y = \tilde{y}\cos\theta + \tilde{x}\sin\theta\,,$$

see Fig.8. The incoming ray is then

$$\tilde{y} = \hat{Y}_I\,,$$

and the two interfaces are

$$\tilde{y} = \tilde{x}\tan\tilde{\alpha}_j + \tilde{\beta}_j \qquad \text{with} \qquad \tilde{\alpha}_j = \alpha_j - \theta\,, \quad \tilde{\beta}_j = \frac{\beta_j\cos\alpha_j}{\cos\tilde{\alpha}_j}\,.$$

The earlier calculations lead to a final ray

$$(\tilde{y} - \tilde{y}_2) = -(\tilde{x} - \tilde{x}_2)\cot(\tilde{\alpha}_2 + \tilde{\varphi}_2)\,,$$

i.e.

$$y = \frac{(\tilde{y}_2 + \tilde{x}_2\cot(\alpha_2 + \tilde{\varphi}_2 - \theta))\sin(\alpha_2 + \tilde{\varphi}_2 - \theta)}{\sin(\alpha_2 + \tilde{\varphi}_2)} - \tilde{x}\cot(\alpha_2 + \tilde{\varphi}_2)\,,$$

with

$$\tilde{\varphi}_2 = \sin^{-1}\left(\frac{n_1}{n_2}\sin(\alpha_1 + \tilde{\varphi}_1 - \alpha_2)\right)\,, \qquad \tilde{\varphi}_1 = \sin^{-1}\left(\frac{n_1}{n_2}\cos(\alpha_1 - \theta)\right)\,,$$

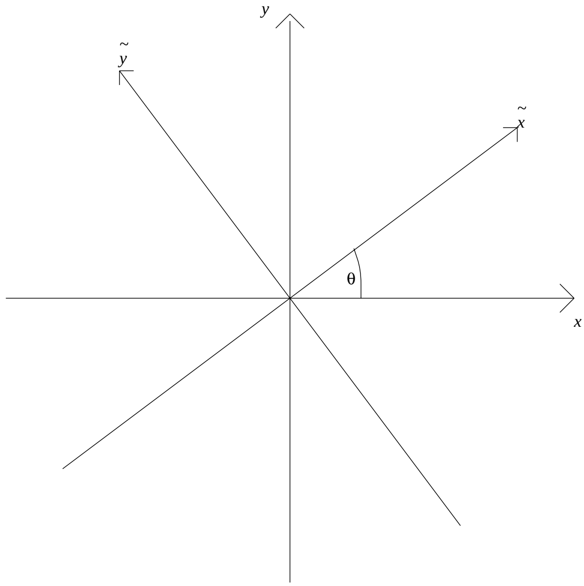

Figure 8: Rotation of axes.

$$\tilde{y}_2 = \tilde{y}_1 - (\tilde{x}_2 - \tilde{x}_1)\cot(\alpha_1 + \tilde{\varphi}_1 - \theta) = \tilde{x}_2 \tan(\alpha_2 - \theta) + \tilde{\beta}_2\,,$$

$$\tilde{y}_1 = \hat{Y}_I \qquad \text{and} \qquad \tilde{x}_1 = (\tilde{y}_1 - \tilde{\beta}_1)\cot(\alpha_1 - \theta)\,.$$

During the workshop, no attempt was made to use these equations to find α's and β's from given transmitted rays. More time was spent on the case of an inclusion with known constant $n_1 = n_0(1+\delta)$, where $\delta \ll 1$.

On returning to (9), and using $n = n_0 + (n_1 - n_0)H(x + (b-y)/a)$ (with H the Heaviside function), we find that dy/dx jumps $-\delta/a$ as a ray, initially parallel to the x axis, crosses the interface between the media $n = n_0$ in $x < (y-b)/a$ and $n = n_1$ in $x > (y-b)/a$. The same linear approximation (9) leads to a second jump in dy/dx, this time of δ/c, as the ray re-enters $n = n_0$ through a surface $y = cx + d$. (These changes may be seen to agree with the above, more general, application of Snell's law, when this limit of small δ is taken: $y = ax + b$ and $y = cx + d$ correspond to $\tan\alpha_1 = a$, $\tan\alpha_2 = b$, then

$$-\left(\frac{\pi}{2} - \alpha_1 - \varphi_1\right) = -\left(\cos^{-1}\left(\frac{n_0}{n_1}\cos\alpha_1\right) - \alpha_1\right) \sim -\cos^{-1}(\cos\alpha_1 - \delta\cos\alpha_1) + \alpha_1$$

$$\sim -\delta\cos\alpha_1/\sqrt{1-\cos^2\alpha_1} = -\delta\cot\alpha_1 = -\delta/a$$

and

$$
\begin{aligned}
-\cot(\alpha_2+\varphi_2) \quad &\sim -\cot\left(\alpha_2+\sin^{-1}\left((1+\delta)\sin\left(\tfrac{\pi}{2}-\tfrac{\delta}{a}-\alpha_2\right)\right)\right) \\
&\sim -\cot\left(\alpha_2+\sin^{-1}\left(\sin\left(\tfrac{\pi}{2}-\alpha_2\right)+\delta\left(\sin\left(\tfrac{\pi}{2}-\alpha_2\right)-\tfrac{1}{a}\cos\left(\tfrac{\pi}{2}-\alpha_2\right)\right)\right)\right) \\
&\sim -\cot\left(\alpha_2+\tfrac{\pi}{2}-\alpha_2+\delta\left(\cos\alpha_2-\tfrac{1}{a}\sin\alpha_2\right)/\sqrt{1-\cos^2\alpha_2}\ \right) \\
&\sim -\cot\left(\tfrac{\pi}{2}+\delta\left(\cot\alpha_2-\tfrac{1}{a}\right)\right) \sim -\tan\left(\delta\left(-\tfrac{1}{c}+\tfrac{1}{a}\right)\right) \sim \left(\tfrac{1}{c}-\tfrac{1}{a}\right).
\end{aligned}
$$

Of course these approximations fail for rays near glancing, *i.e.* a or c small, when angles of deviation become significant.)

A ray initially given by $y = Y_I$ meets $y = ax + b$ at $x_1 = (Y_I - b)/a$, then follows the path $y \sim Y_I - \delta(x - (Y_I - b)/a)/a$, reaches $y = cx + d$ at $x_2 \sim (Y_I - d)/c$, $y_2 \sim Y_I + \delta((c-a)Y_I - bc + ad)/a^2c$, and then departs along

$$
y \sim Y_I + \delta\left(\left(\frac{1}{c}-\frac{1}{a}\right)x + \left(\frac{1}{a^2}-\frac{1}{c^2}\right)Y_I + \left(\frac{d}{c^2}-\frac{b}{a^2}\right)\right). \tag{18}
$$

Changing the direction of the laser by an angle θ, and using new co-ordinates $\tilde{x}$, $\tilde{y}$ so that the $\tilde{x}$ axis is parallel to the beam, is equivalent to a rotation of the pre-form or fibre through an angle θ in the clockwise direction (as above). For a simple method (meaning that we wanted linear problem a if possible), the angle θ was assumed to be small; then

$$
x \sim \tilde{x} - \theta\tilde{y}\,, \qquad y \sim \tilde{y} + \theta\tilde{x}\,.
$$

The two interfaces can then be written as

$$
\tilde{y} \sim (a - (1+a^2)\theta)\tilde{x} + b(1 - a\theta)\,, \quad \tilde{y} \sim (c - (1+c^2)\theta)\tilde{x} + d(1 - c\theta)\,.
$$

If the laser beam is essentially as before, $\tilde{y} = Y_I$, the final beam is now

$$
\begin{aligned}
\tilde{y} \quad &\sim Y_I + \delta\left[\left(\left(\frac{1}{c}+\frac{1+c^2}{c^2}\theta\right)-\left(\frac{1}{a}+\frac{1+a^2}{a^2}\theta\right)\right)\tilde{x} + \left(\frac{1}{a^2}+\frac{2(1+a^2)}{a^3}\theta-\frac{1}{c^2}+\frac{2(1+c^2)}{c^3}\theta\right)Y_I\right. \\
&\quad \left. + \frac{d}{c^2} + \left(\frac{2d}{c^3}(1+c^2)-\frac{d}{c}\right)\theta - \frac{b}{a^2} - \left(\frac{2b}{a^3}(1+a^2)-\frac{b}{a}\right)\theta\right] \\
&\sim Y_I + \delta\left[\left(\left(\frac{1}{c}-\frac{1}{a}\right)+\left(\frac{1}{c^2}-\frac{1}{a^2}\right)\theta\right)\tilde{x} + \left(\left(\frac{1}{a^2}-\frac{1}{c^2}\right)Y_I + \left(\frac{d}{c^2}-\frac{b}{a^2}\right)\right)\right. \\
&\quad \left. + \left(2\left(\frac{1+a^2}{a^3}-\frac{1+c^2}{c^3}\right)Y_I + \left(\frac{d(2+c^2)}{c^3}-\frac{b(2+a^2)}{a^3}\right)\right)\theta\right],
\end{aligned}
$$

neglecting terms in δ^2 and in θ^2 (*i.e.* we retain "constant terms" down to size δ and the largest terms involving θ).

Both final beams take the same form:

$$
y \sim Y_I + \delta(A_0 x + B_0) \quad \text{and} \quad \tilde{y} \sim Y_I + \delta(A_s\tilde{x} + B_s)
$$

where the directions A_0 and A_s,

$$A_0 = \frac{1}{c} - \frac{1}{a} \quad \text{and} \quad A^* \equiv \frac{A_s - A_0}{\theta} = \frac{1}{c^2} - \frac{1}{a^2}\,, \tag{19}$$

and the offsets B_0 and B_s,

$$B_0 = \left(\frac{1}{a^2} - \frac{1}{c^2}\right) Y_I + \left(\frac{d}{c^2} - \frac{b}{a^2}\right) \tag{20}$$

and

$$B^* \equiv B_s - B_0 = 2\left(\frac{1+a^2}{a^3} - \frac{1+c^2}{c^3}\right) Y_I + \left(\frac{d(2+c^2)}{c^3} - \frac{b(2+a^2)}{a^3}\right), \tag{21}$$

are, in principle, measureable. (Of course these measurements need to be made with an accuracy better than $O(\delta\theta)$. It should be emphasized that this method works only if δ is known; otherwise at least one more measurement is needed.)

Given A_0 and A_s, (19) can be solved to give

$$a = -2A_0/(A_0^2 + A^*) \quad \text{and} \quad c = a/(A_0 + 1)\,.$$

The values of b and d are then obtained from the linear equations (20) and (21):

$$b = \frac{Y_I(2 + a^2 - ac)}{(2 - ac)} + \frac{a^3 c(B^* - B_0(c + 2/c))}{(a - c)(2 - ac)}$$

and

$$d = c^2 B_0 + c^2 b/a^2 + (1 - c^2/a^2) Y_I\,.$$

It can be seen that b and d cannot be determined if $a = c$, *i.e.* if the planes are parallel. This is to be expected as a translation of an infinite slab of material of $n = n_1$ always leaves the outgoing ray unaltered. (There is also a problem with $ac = 2$, $\tan\alpha_1 \tan\alpha_2 = 2$ in the earlier notation. It is not immediately clear what the reason for this is.) Poor results in obtaining the position, not the slopes, will occur if the planes are close to being parallel. This was confirmed in some numerical experiments doing these computations for rays found by solving (6) with a smoothed n. The results for $a \not\approx c$ were much better.

Given that co-ordinates are better fixed with respect to the object under investigation, the final ray may instead be written using the co-ordinates x, y:

$$\begin{aligned} y \sim\ & Y_I + \left(\theta + \left(\frac{1}{c} - \frac{1}{a}\right)\delta + \left(\frac{1}{c^2} - \frac{1}{a^2}\right)\delta\theta\right) x \\ +\ & \delta\left[\left(\left(\frac{1}{c^2} - \frac{1}{a^2}\right) Y_I + \left(\frac{d}{c^2} - \frac{b}{a^2}\right)\right) + \left((Y_I - b)\left(\frac{2}{a^3} + \frac{1}{a}\right) - (Y_I - d)\left(\frac{2}{c^3} + \frac{1}{c}\right)\right)\theta\right]. \end{aligned} \tag{22}$$

Of course angles θ which are not small can be employed: this would have the advantage of allowing less precise measurements to be used. There are, however, two difficulties. First, the equations for the slopes now turn out to be equivalent to a quadratic equation, so the correct root has to be chosen. The second is that there is the possibility that the points of entry to and exit from the region Ω (where $n = n_1$) can move significantly so, except for this slab geometry, the slopes a and c can also be expected to change considerably.

For a more general region Ω, with a curved boundary $\partial\Omega$, there is an added complication due to the curvature of $\partial\Omega$: the differing slopes of $\partial\Omega$ at the different points at which the rays enter and exit Ω affect the paths. To get round this difficulty, a third ray or beam, $y = Y_I + Y_T\theta$, can be employed. This results in another two measured quantities, the new offset and direction, and provides information for the two extra unknowns: the curvatures, or second derivatives, at the points the rays cross $\partial\Omega$. (The linearity inherent in this basic technique means that there is no point in using other small angles or translations, or employing the two perturbations simultaneously.)

For simplicity, the boundary $\partial\Omega$ is written, locally, as $y = f(x)$. We continue to assume that the rays cross $\partial\Omega$ only twice. (This is certain if Ω is convex and can always be arranged if Ω is star-shaped – although some prior knowledge of the shape would be needed. It might be possible to cope with more general regions if we can apply this basic method to find parts of $\partial\Omega$, and then use this knowledge to allow for some deviations of ray path in trying to determine other parts.)

We now take $x_{10} < x_{20}$ to be the values of x where the first ray, initially $y = Y_I$, crosses $y = f(x)$:

$$Y_I = f(x_{10})\,; \qquad Y_I \sim f(x_{20})$$

and more specifically,

$$x_{20} \sim x_2 + (x_1 - x_2)/f'(x_1)f'(x_2) \quad \text{ with } \ f(x_2) = Y_I\,.$$

Also $f'(x_1)$, $f'(x_2)$ replace a, c above: the first ray leaves along

$$y - Y_I \sim \delta\left(\left(\frac{1}{f'(x_2)} - \frac{1}{f'(x_2)}\right)x + \left(\frac{x_1}{f'(x_1)} - \frac{x_2}{f'(x_2)}\right)\right),$$

i.e.

$$A_0 = 1/f'(x_2) - 1/f'(x_1)\,, \qquad B_0 = x_1/f'(x_1) - x_2/f'(x_2)\,.$$

Consideration of the rotated case, with careful calculation of just where the ray crosses $\partial\Omega$, leads eventually to

$$y - Y_I \sim \theta x + \delta(A_R x + B_R)$$

with

$$A_R = A_0 + \theta A_R^*\,, \quad B_R = B_0 + \theta B_R^*\,,$$
$$A_R^* = 1/f'(x_2)^2 - 1/f'(x_1)^2 + x_1 f''(x_1)/f'(x_1)^3 - x_2 f''(x_2)/f'(x_2)^3\,,$$
$$B_R^* = x_2^2 f''(x_2)/f'(x_2)^3 - x_1^2 f''(x_1)/f'(x_1)^3 + 2x_1/f'(x_1)^2 - 2x_2/f'(x_2)^2 + x_1 - x_2\,.$$

(This can be seen to agree with (22) when $f''(x_1) = f''(x_2) = 0$.) In a similar manner, the translated case gives a final ray

$$y - Y_I \sim Y_T\theta + \delta(A_T x + B_T)$$

with

$$A_T = A_0 + Y_T\theta A_T^*\,, \quad B_T = B_0 + Y_T\theta B_T^*\,, \quad A_T^* = f''(x_2)/f'(x_2)^3 - f''(x_1)/f'(x_1)^3\,,$$

$$B_T^* = x_1 f''(x_1)/f'(x_1)^3 - x_2 f''(x_2)/f'(x_2)^3 + 1/f'(x_2)^2 - 1/f'(x_1)^2\,.$$

The basic idea for using these equations would be to regard A_0, B_0, A_R^*, B_R^*, A_T^*, B_T^* as known from the measurements. In line with the polygonal case, we may write $f'(x_1) = a$, $f(x_1) = ax_1 + b =$

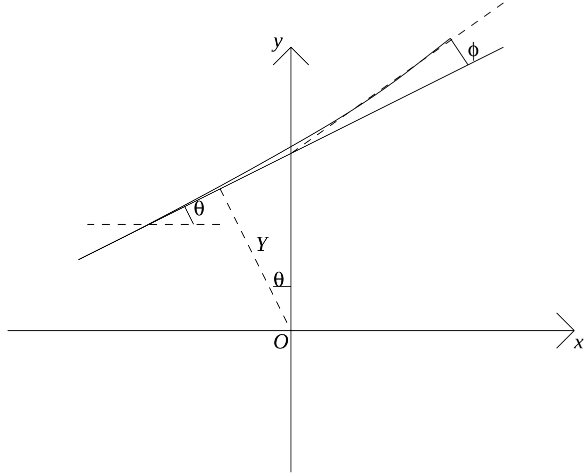

Figure 9: A general deflected beam.

Y_I, $f'(x_2) = c$, $f(x_2) = cx_2 + d = Y_I$ so we wish to find a, b, c, d and hence $x_1 = (Y_I - b)/a$ and $x_2 = (Y_I - d)/c$. The unknowns, on making these substitutions, are then a, b, c, d, $f''(x_1)$ and $f''(x_2)$. The last two can be found, in terms of the other four, from, for example, A_T^* and B_T^*. This then allows, at least in principle, the remaining values to be obtained from the other four measurements.

1.4.3 Circular symmetry.

The work of 1995 dealt with radial symmetry as long as the refractive index only varies slightly. The solution (12) might be used more genrally in the following way.

With reference to an origin O, a two-parameter family of measurements can be made:

$\phi(Y, \theta) =$ the angle of deflection for a beam initially making an angle θ with the x axis and directed to pass Y (to the left) of O; see Fig.9.

From the linearity of this problem, taking the average over θ, for fixed Y, gives the deflection produced by a symmetric refractive index:

$$\bar{\phi}(Y) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(Y, \theta)\, d\theta = \text{ deflection resulting from } \bar{n}(r)\,,$$

$$\bar{n}(r) = \frac{1}{2\pi}\int_{-\pi}^{\pi} n(r\cos\theta, r\sin\theta)\, d\theta\,.$$

Then, from (12),

$$\bar{n}(r) - n_0 = -\frac{n_0}{\pi}\int_r^{\infty} \frac{\bar{\phi}(Y)\, dY}{(Y^2 - r^2)^{1/2}}.$$

In particular,

$$n(O) = \bar{n}(0) \quad = n_0 - \frac{n_0}{\pi} \int_0^\infty \frac{\bar{\phi}(Y)}{Y}\, dY$$

$$= n_0 - \frac{n_0}{2\pi^2} \int_0^\infty \int_{-\pi}^{\pi} \frac{\phi(Y,\theta)}{Y}\, d\theta\, dY$$

$$= n_0 - \frac{n_0}{2\pi^2} \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{\varphi(x,y)}{x^2+y^2}\, dx\, dy$$

with $\varphi(x,y) =$ deflection (to the left) of a ray which, if it had not been deflected, would have had (x,y) as its closest point to O: $x = -Y\sin\theta$, $y = Y\cos\theta$. More generally,

$$n(x_o,y_o) = n_0 - \frac{n_0}{2\pi^2} \int_{-\infty}^\infty \int_{-\infty}^\infty \frac{\Phi(x,y;x_o,y_o)}{(x-x_o)^2+(y-y_o)^2}\, dx\, dy$$

where Φ is the deflection of a beam originally directed to have (x,y) as its closest point to (x_o,y_o).

One further comment might be made on this "linear", symmetric case and how it relates to C.A.T. (computer aided tomography). The absorption of a ray passing straight through a circularly symmetric medium with (linear) absorption coefficient $\sigma(r)$ is given by

$$\alpha(Y) = \int_{-\infty}^\infty \sigma(\sqrt{x^2+y^2})\, dx = \int_Y^\infty \frac{r\sigma(r)\, dr}{\sqrt{r^2-Y^2}}\,,$$

for the ray $y \equiv Y$. The determination of $r\sigma(r)$, and hence of σ, now follows from solving an equation of the form (12) to get an answer like (11) – which contains a derivative of the measured quantity. This suggests that C.A.T. is not so well posed as the one under investigation here.

The method applied to the radially symmetric problems with $n \approx n_0$ can be extended to those with n varying significantly (but now the lack of linearity for the geometric-optics formulation means that there is no apparent way of applying the results to asymmetric cases). Now we return to (7) and (8), from which it is seen that for a ray coming from $x = -\infty$ along $y = Y_I$, *i.e.*

$$r \to \infty\,, \qquad x \to -\infty\,, \qquad y \to Y_I\,, \qquad \theta \to \pi \quad \text{as } s \to -\infty\,,$$

$$\frac{dr}{ds} = -1 + O(1/r^2) \qquad \text{so} \qquad r \sim -s\,.$$

Also $n(r) \to n_0$ as $r \to \infty$ which means that $d\theta/ds \sim -Y/s^2$ and $\theta \sim \pi + Y/s$.

Then $y = r\sin\theta \sim (-s) \times (-Y/s)$ and Y is precisely Y_I.

The arc length s can be eliminated from (7) and (8):

$$\frac{dr}{d\theta} = \pm\frac{\sqrt{1 - n_0^2Y^2/n^2r^2}}{n_0Y/nr^2} = \pm\frac{r}{Y}\sqrt{n_r^2r^2 - Y^2}\,, \tag{23}$$

where $n_r(r) = n(r)/n_0$ and, assuming that $Y > 0$, the positive sign applies as the light approaches and the negative sign when it departs (θ decreases along the whole of the ray, r decreases then increases). It can be seen from (23) that the minimum value of r, r_m, taken when $dr/d\theta = 0$,

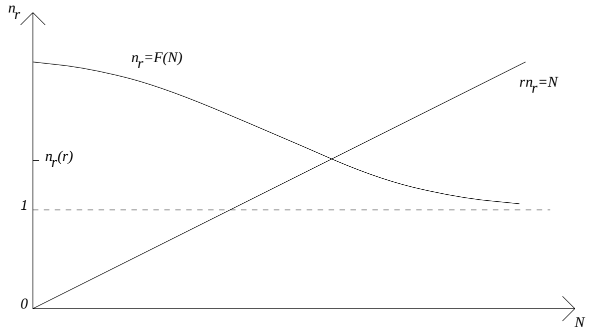

Figure 10: Finding n_r from F.

satisfies $r_m n_r(r_m) = Y$. As r decreases from ∞ to r_m and then increases to infinity again, θ decreases from π to $\phi(Y)$, with, as usual, $\phi(Y)$ being the angle of deflection (measured to the left) for a ray directed along $y = Y$ (> 0) from $x = -\infty$. Since

$$\frac{d\theta}{dr} = \pm \frac{Y}{r\sqrt{n_r^2 r^2 - Y^2}} ,$$

$$\pi - \phi(Y) = 2Y \int_{r_m}^{\infty} \frac{dr}{r\sqrt{n_r(r)^2 r^2 - Y^2}} .$$

(Elementary integration provides the check that for a uniform medium, $n_r \equiv 1$, there is no refraction:

$$Y \int_Y^{\infty} \frac{dr}{r\sqrt{r^2 - Y^2}} \equiv \frac{\pi}{2} .) \tag{24}$$

At this stage, a change of variables is appropriate: writing $N = rn_r(r)$, so that $N = Y$ at $r = r_m$,

$$\pi - \phi(Y) = 2Y \int_Y^{\infty} \left(\frac{1}{r}\frac{dr}{dN}\right) \frac{dN}{\sqrt{N^2 - Y^2}} = 2Y \int_Y^{\infty} \frac{\frac{dR}{dN} dN}{\sqrt{N^2 - Y^2}} ,$$

on writing $R = \ln r$. As $r \to \infty$, $n_r \to 1$, $N \sim r$ and $dR/dN \sim 1/N$. The above identity, (24), provides the slightly changed integral equation

$$\phi(Y) = -2Y \int_Y^{\infty} \frac{\frac{d}{dN}(R - \ln N)}{\sqrt{N^2 - Y^2}} dN = 2Y \int_Y^{\infty} \frac{\frac{d}{dN} \ln n_r}{\sqrt{N^2 - Y^2}} dN.$$

This <u>exact</u> Abel integral equation (which reduces to (11) for $n_r = n/n_0 = 1 + \hat{n}$ with $|\hat{n}(r)| \ll 1$) has the solution

$$n_r = \exp\left\{-\frac{1}{\pi}\int_N^{\infty} \frac{\phi(Y)\, dY}{\sqrt{N^2 - Y^2}}\right\} \equiv F(N)$$

so n_r can be found as a function of N. The final stage of determining $n(r)$ is given by observing that for a given r, $n_r(r) = n(r)/n_0$ is the solution of $n_r = F(rn_r)$ (see Fig.10).

2 Designing a Grating

2.1 Introduction

The synthesis of gratings problem was the second out of three problems presented by Corning in the area of "inverse problems for optical fiber device measurement and design". Interest in non-uniform grating devices stems from the possibilities for manufacturing grating devices such as wavelength-selective mirrors, wavelength-selective couplers, filters, *etc.* The intention is to design special non-uniform gratings that will have some specified desired transmission and reflection properties. This leads to mathematical inverse problems of computing the medium properties (refractive index) that, when solving Maxwell's equations for light propagation in that medium, will result in a solution with desired properties.

An alternative approach, but one only pertaining to desired reflection, is considered in §2.7.

2.2 Mathematical formulation

We consider an idealized field propagating in one dimension, z, along the grating structure. The electric and magnetic fields in the grating structure can be written as

$$\begin{aligned} E(z,t) &= \hat{x}\,\mathrm{Re}\{E(z)\exp(-i\omega t)\}, \\ H(z,t) &= \hat{y}\,\mathrm{Re}\{H(z)\exp(-i\omega t)\}. \end{aligned} \tag{25}$$

We define a new spatial variable, $\xi = k_o z$, where k_0 is the light wavenumber at the Bragg scattering resonance (corresponding to the wavelength $\lambda_0 \approx 1.5 \times 10^{-6}\,\mathrm{m}$). The new variable ξ leads to a better scaling of the equations which may be important for the numerical solution (for $\xi = 1$, $z \approx \lambda_0/2\pi$).

Substitution of (25) into Maxwell's equations and following the derivation of [16], the following approximation is obtained:

$$E(\xi) = u(\xi)\exp\left(\frac{i}{2}\phi(\xi)\right)\exp(i\xi) + v(\xi)\exp\left(-\frac{i}{2}\phi(\xi)\right)\exp(-i\xi), \tag{26}$$

where $u(\xi)$ and $v(\xi)$ correspond to the components of the light propagating forwards and backwards respectively (in the directions of $+z$ (transmitted) and $-z$ (reflected), respectively) and satisfy the following "state" equations,

$$\frac{dU}{d\xi} = M(\xi, n(\xi))U, \tag{27}$$

where $U(\xi) = (u(\xi), v(\xi))^T$ and $M(\xi, n(\xi))$ is a 2×2 matrix that depends on the spatial variable, ξ, and on the refractive index, $n(\xi)$ (see (33) below). Note that stationary equations are obtained as a result of extracting the term $e^{i\omega t}$ from Maxwell's equations. Also note that in terms of the spatial variable z, the left-hand side of equation (27) would have been about six orders of magnitude smaller than the right-hand side, which is undesirable for the numerical solution.

The boundary conditions are determined by specifing the forward propagating light at the inlet of the grating, $\xi = 0$, to be the input light, u_0, and the condition that there are no backward traveling waves at the outlet of the grating, $\xi = \xi_M$:

$$u(0,\omega) = u_0\,, \qquad v(\xi_M,\omega) = 0\,. \tag{28}$$

Given the grating refractive index, $n(\xi)$, the state equation is a well posed boundary value problem for a system of ODEs.

2.3 The inverse problem

We search for a refraction index, $n(\xi)$, such that $u(\xi_N)$ and $v(\xi_0)$ will be **equal** to some desired values, for a range of incoming light frequencies, ω.

We conjecture that such a refraction index may not exist for arbitrary $u(\xi_N, \omega)$ and $v(0, \omega)$ over a range of incoming light frequencies, ω. We therefore suggest that the problem be posed as an optimization problem rather than as an inverse problem, *i.e.*, find the refraction index, $n(\xi)$, such that $u(\xi_N, \omega)$ and $v(0, \omega)$ will be **as close as possible** to the desired values, for a range of incoming light frequencies, ω.

2.3.1 Low-dimensional model for the refraction index.

Another difficulty with the above formulation is the excess of freedom in the choice of the refraction index, $n(\xi)$, leading to an underdetermined problem. For example, certain translations in the refraction index, $n(\xi)$, lead to the same transmission and reflection properties (*c.f.* the difficulty in locating parallel surfaces in §1). That difficulty can be solved by reducing the refraction index to a lower-dimensional model. Instead of searching for any index function, $n(\xi)$, we choose to consider only indices of the following form, [16],

$$n(\xi) = n_0(1 + \sigma(\xi) + 2\kappa(\xi)\cos[2\xi + \phi(\xi)]), \tag{29}$$

where $\sigma(\xi)$ is the "average" index, κ is the envelope of the grating's variation, also known as the "appodization" function, and ϕ is the "chirp" function (determining the grating wavelength). (See Fig.11 for a sketch of such a refractive-index profile.)

We define the six parameters, $\alpha_1, \cdots, \alpha_6$, as the "design" parameters of the refraction index in (29):

$$\begin{aligned} \sigma(\xi) &= \alpha_1 ; \\ \phi(\xi) &= \alpha_2\xi + \alpha_3\xi^2 ; \\ \kappa(\xi) &= \alpha_6 \exp(-\alpha_4 \xi^{2\alpha_5}) . \end{aligned} \tag{30}$$

Eqs.(29) and (30) reduce the one-dimensional function space (that the function $n(\xi)$ belongs to) into a six-dimensional vector space. That space will be refered to as the "design space" from here on.

2.3.2 The state equations.

In terms of the model (29), the state equations (27) are given by

$$\begin{aligned} \frac{du(\xi)}{d\xi} &= i[\hat{\sigma}(\xi, \alpha_1, \alpha_2, \alpha_3)u(\xi) + \kappa(\xi, \alpha_4, \alpha_5, \alpha_6)v(\xi)] , \\ \frac{dv(\xi)}{d\xi} &= -i[\hat{\sigma}(\xi, \alpha_1, \alpha_2, \alpha_3)v(\xi) + \kappa(\xi, \alpha_4, \alpha_5, \alpha_6)u(\xi)] , \end{aligned} \tag{31}$$

with the boundary conditions (28), where $\hat{\sigma}(\xi)$ and the "detuning" parameter Δ given by

$$\begin{aligned} \hat{\sigma}(\xi) &= \sigma(\xi) + \Delta - \frac{1}{2}\frac{\partial \phi}{\partial \xi}, \\ \Delta &= \frac{\omega - \omega_0}{\omega_0}, \end{aligned} \tag{32}$$

where $\omega_0 = ck_0/n_0$.

(The matrix M used above is given by

$$M = \begin{pmatrix} i\hat{\sigma} & i\kappa \\ -i\kappa & -i\hat{\sigma} \end{pmatrix}, \tag{33}$$

again see [16].)

2.4 Optimal control approach

In the optimal-control formulation of the problem, we make a distinction between state variables (u, v) and control (or design) variables α. A multi-objective cost functional F is defined in terms of the state and control variables. The problem is then to find the set of control variables that minimize the cost functional with the state equations acting as constraints: $\min_{\alpha,u,v} F(u, v)$, subject to the state equations $h(\alpha, u, v) = 0$.

2.4.1 The optimization problem.

We must solve

$$\min_{\alpha_1,\cdots,\alpha_6,u,v} F(u,v) = \Big[\sum_{\omega} \Big(|u(\xi_M,\omega)| - T^d(\omega) \Big)^2 + \mu \sum_{\omega} \Big(|G(u(\xi_M,\omega))| - \phi^d(\omega) \Big)^2 \Big], \tag{34}$$

subject to the state equations (31), (28). Here, a discrete set of frequencies are used in the optimization.

The first term in the cost functional, (34), models the difference of the transmitted light, at the right end of the grating, from a desired transmission, $\phi^d(\omega)$. The second term models the difference of the phase at the right end of the grating from a desired phase, $\phi^d(\omega)$, and is given by

$$G(u(\xi_M,\omega)) = \tan^{-1} \frac{\mathrm{Im}\{u(\xi_M,\omega)\}}{\mathrm{Re}\{u(\xi_M,\omega)\}}.$$

2.4.2 The necessary conditions for a minimum.

The co-state equations are given by

$$\begin{aligned} \frac{d\Lambda_u}{d\xi}(\xi) &= i[\hat{\sigma}(\xi,\alpha_1,\alpha_2,\alpha_3)\Lambda_u(\xi) - \kappa(\xi,\alpha_4,\alpha_5,\alpha_6)\Lambda_v(\xi)], \\ \frac{d\Lambda_v}{d\xi}(\xi) &= -i[\hat{\sigma}(\xi,\alpha_1,\alpha_2,\alpha_3)\Lambda_v(\xi) - \kappa(\xi,\alpha_4,\alpha_5,\alpha_6)\Lambda_u(\xi)], \end{aligned} \tag{35}$$

with the boundary conditions,

$$\Lambda_u(\xi_M,\omega) = F_u(u,v)\,, \quad \Lambda_v(0,\omega) = 0\,, \tag{36}$$

where

$$\begin{aligned} F_u(u,v) \;&=\; 2\Big[\sum_\omega \Big(|u(\xi_M,\omega)| - T^d(\omega)\Big)\frac{d}{du}|u(\xi_M,\omega)| \\ &+\; \mu\sum_\omega \Big(|G(u(\xi_M,\omega))| - \phi^d(\omega)\Big)\frac{d}{du}|G(u(\xi_M,\omega))|\Big]. \end{aligned} \tag{37}$$

The derivative

$$\frac{d}{du}|a+ib| = \left(a\frac{da}{du} + b\frac{db}{du}\right)/|a+ib|$$

is used in evaluating the right-hand side of (37).

Note that the boundary conditions (36) are complementary to those of the state equation, (28), *i.e.*, the adjoint forward propagating wave is specified at the outlet of the grating, while the adjoint backward propagating wave is specified at the inlet of the grating.

The optimality condition is given by

$$-\int_0^{\xi_M} d\xi \Lambda^T \frac{d}{d\alpha}\begin{pmatrix} \hat{\sigma} & \kappa \\ \kappa & \hat{\sigma} \end{pmatrix} U = 0\,, \tag{38}$$

where $U = (u,v)^T$ and $\Lambda = (\Lambda_u, \Lambda_v)^T$.

The sensitivity gradient, $\nabla_\alpha F$, is given by the residual of the optimality condition (38), for states and co-states that satisfy Eqs.(31) and (35) respectively.

2.5 The algorithm

1. Start with an initial guess of the solution $\alpha_1, \cdots, \alpha_6$.
2. Compute $\hat{\sigma}(\alpha_1,\alpha_2,\alpha_3)$ and $\kappa(\alpha_4,\alpha_5,\alpha_6)$
3. Solve the state equation (31) for a range of light frequencies, e.g. using Runge-Kutta shooting-Newton method, and obtain U.
4. Compute the cost functional (34), $F(\alpha)$.
5. Solve the co-state equation (35) for a range of light frequencies, e.g. using Runge-Kutta shooting-Newton method, and obtain Λ.
6. Integrate (38) with the trapezoidal method and obtain the sensitivity gradient, $\nabla_\alpha\ F$.
7. Update the design variables, with the gradient information, e.g. by a Quasi-Newton trust-region method.
8. Unless the gradient is smaller than some tolerance value ϵ,

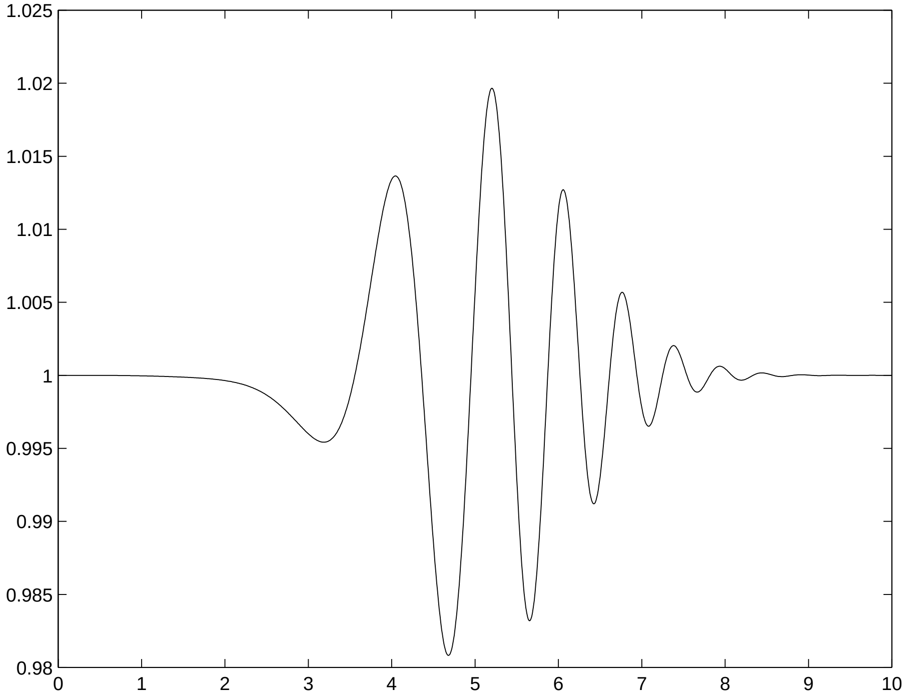

Figure 11: The refractive index for $(\vec{\alpha} = (0, 2, -1, 10, 1, 10^{-2})$, and $0 \leq \xi \leq 10$, $\lambda_0 = 1.5\,\mu$m.

$|\nabla_\alpha\ F|_2 < \epsilon$, `go back to step number 3.`

Note that some optimization packages do not require the user to provide the gradient, only the cost functional, so steps number $4-6$ can be skipped at the expense of a more computationally intensive algorithm.

2.6 Preliminary numerical results

During the workshop, steps $1-6$ were implemented numerically using Matlab. There was insufficient time to include the optimization process for the full solution of the problem.

Figs. 12 and 13 depict the real and imaginary solutions of the state equation. Note that $v(\xi)$ is equal to zero at the outlet.

Figs. 14 and 15 depict the real and imaginary solutions of the co-state equation. Note that $\Lambda_v(\xi)$ is equal to zero at the inlet.

The light frequency in the simulations is taken to be $\omega = 1.22\omega_0$, where $\omega_0 = 2\pi/\lambda_0$ is the resonance frequency.

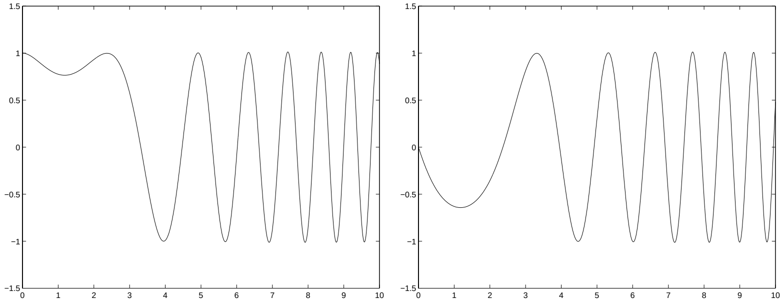

Figure 12: Real and imaginary parts of the right-propagating wave, $u(\xi)$.

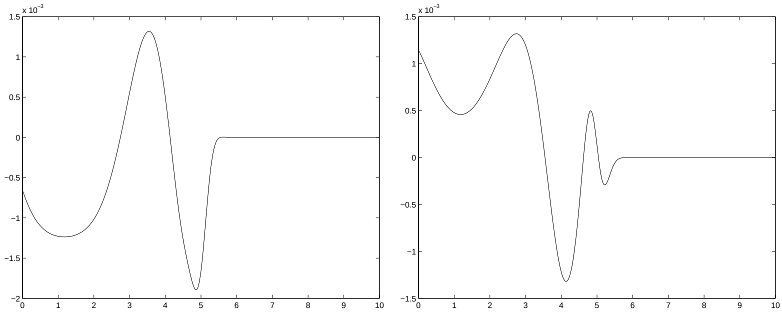

Figure 13: Real and imaginary parts of the left-propagating wave, $v(\xi)$.

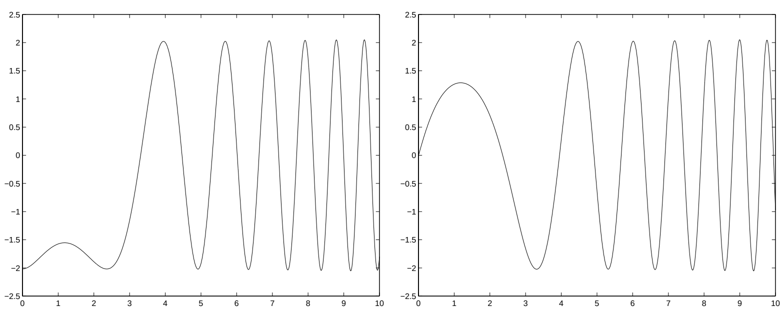

Figure 14: Real and imaginary parts of the **adjoint** right-propagating wave, $\Lambda_u(\xi)$.

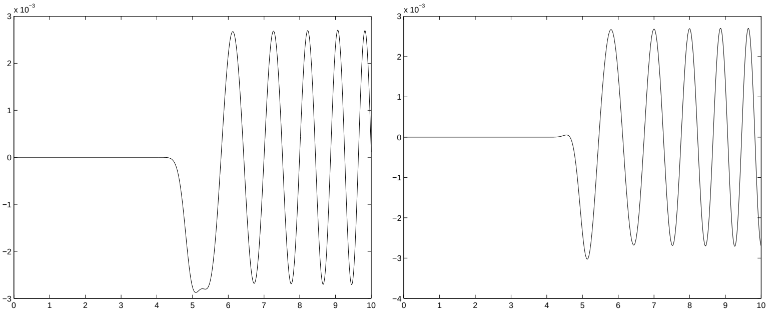

Figure 15: Real and imaginary parts of the left **adjoint** propagating wave, $\Lambda_v(\xi)$.

2.7 A time-domain approach

In this section, rather than attempt to obtain a grating which has desired transmitted and reflected waves for a harmonic incident wave, the problem of getting a specified (transient) reflected signal from a given transient input is considered. Following [10] (but modifying it a little), we model the problem as follows. The grating is one-dimensional and z measures the distance from the left end of the grating. A complex-valued function, $q(z)$ on $(-\infty, \infty)$ characterizes the grating and we have $q(z) = 0$ for $z \leq 0$. If $l(z,t)$ and $r(z,t)$ are the "left" and "right" going waves then

$$\frac{\partial l}{\partial z} - \frac{\partial l}{\partial t} = q(z) r, \qquad \forall\, t,\ \forall\, z, \tag{39}$$

$$\frac{\partial r}{\partial z} + \frac{\partial r}{\partial t} = \overline{q}(z) l, \qquad \forall\, t,\ \forall\, z. \tag{40}$$

Here $\overline{q}$ is the complex conjugate of q. We have the initial conditions

$$l(z,t) = 0, \qquad r(z,t) = f(t-z), \qquad \text{for } t < 0\ , \tag{41}$$

with $f(t)$ zero for $t < 0$. So the input is is a right-moving wave coming from the left of the grating – there is no wave coming from the right end of the grating. Note that because $q(z)$ is zero for $z \leq 0$ and $f(t)$ is zero for $t < 0$, the initial conditions satisfy the PDE.

Since $q(z)$ is zero for $z \leq 0$ we have $\partial l/\partial z - \partial l/\partial t = 0$ in the region $z \leq 0$ for all t. Hence in the region $z \leq 0$, $l(z,t) = g(t+z)$ for some function $g(t)$; this vanishes for $t < 0$ because $l(z,t)$ is zero for $t < 0$. Then $g(t+z)$ is the reflection of the input $f(t-z)$ by the grating. The goal here is simply to choose $q(z)$ so that for a given $f(t)$, $g(t)$ has the desired characteristics.

Since the solution of (39) - (41) is linear in f, we have that $g(t+z) = l(z,t) = L(z,t) * f(t)$ where $L(z,t), R(z,t)$ are the Green's function of the problem, that is, they are solutions of the initial value problem

$$\frac{\partial L}{\partial z} - \frac{\partial L}{\partial t} = q(z) R, \qquad \forall\, t,\ \forall\, z\,, \tag{42}$$

$$\frac{\partial R}{\partial z} + \frac{\partial R}{\partial t} = \overline{q}(z) L, \qquad \forall\, t,\ \forall\, z\,, \tag{43}$$

with the initial conditions

$$L(z,t) = 0, \qquad R(z,t) = \delta(t-z), \qquad \text{for } t < 0 \; . \tag{44}$$

($*$ denotes a convolution: $L(z,t) * f(t) = \int_{-\infty}^{\infty} L(z,t-s) f(s)\, ds$.)

As before, in the region $z \le 0$, $L(z,t) = G(t+z)$ for some $G(t)$ supported in $[0,\infty]$. For $z \le 0$, we have

$$g(t+z) = l(z,t) = L(z,t) * f(t) = G(t+z) * f(t)$$

and hence $g(t) = G(t) * f(t)$. So for $g(t)$ to have the desired characteristics we will need to find a q so that $G(t)$ has certain properties. Or we may rephrase the problem in the following manner:

> *Given a function $G(t)$ (the desired output profile) supported in $[0,\infty)$ determine a function $q(z)$ supported in $[0,\infty)$ so that $L(z,t) = G(t+z)$ for $z \le 0$*

(assuming that we can calculate G from a given f and a desired g).

This leads to several questions which must be resolved. Consider the forward map relating the grating design q to the reflected signal $G(.)$:

$$\mathcal{F} \; : \; q \to G(.) \; .$$

Then we must answer the following questions

1. What is the domain of $\mathcal{F}$, *i.e.* what class of q do we allow as possible grating designs?

2. What is the range of $\mathcal{F}$, *i.e.* what are the possible profiles for the reflected signal and is the desired profile in the range of $\mathcal{F}$?

3. Is $\mathcal{F}$ injective? That is, could there be more than one design q generating the same reflected signal?

4. How do we construct the inverse of $\mathcal{F}$? Given a $G(.)$ in the range of $\mathcal{F}$ how do we recover the q so that $\mathcal{F}(q) = G(.)$.

5. Is $\mathcal{F}$ continuous and is its inverse continuous (assuming $\mathcal{F}$ is injective)? The continuity of the inverse of $\mathcal{F}$ would imply that small errors in the data $G(.)$ would not result in dramatically different grating designs q.

Note that $\mathcal{F}$ is a nonlinear map because the solution of the PDE (42) - (44) depends nonlinearly on q. Hence answering these questions is nontrivial.

Concerning the domain of $\mathcal{F}$, one of the desired features is likely to be that the grating be of a finite length, so $q(z)$ is supported over a finite interval, say $[0, X]$. We will assume this for all our work below.

2.7.1 Real-valued q.

If we only allow real-valued q (which of course is not the problem of interest) these questions have been completely resolved. In this case, the domain of $\mathcal{F}$ is the class of square-integrable functions on $[0, X]$. Further, knowing $G(t) = L(z=0,t)$ for $0 \le t \le 2X$ completely determines q. So $G(t) = L(0,t)$ on $0 \le t \le 2X$ completely determines $G(t) = L(0,t)$ for all $t > 2X$ and hence only the values of $G(t)$ over the interval $[0, 2X]$ may be specified somewhat arbitrarily (they have to satisfy an additional condition stated below). So we have

$$\mathcal{F} \; : \; L^2[0, X] \to L^2[0, 2X]$$

(both q and G are square integrable) and

$$\mathcal{F}(q) = L(z = 0 \, , .) \; .$$

This map is injective, continuous, and its inverse is also continuous. Its range consists exactly of all real-valued functions $G(.)$ in $L^2[0, 2X]$ so that the linear operator

$$\begin{array}{rcl} L^2[0, 2X] & \to & L^2[0, 2X] \\ f & \mapsto & G * f \end{array}$$

has norm strictly less than 1 ($| \int_0^{2X} (G * f)(t) \, dt| < C|f|$, with $C < 1$ and independent of f). This last condition on G is the assertion that the energy of the reflected wave should be uniformly (in f) smaller than the energy of the incident wave. The norm of the above linear operator is some quantity less than or equal to $\|G\|_1$ - the L^1 norm of G.

The injectivity, continuity of $\mathcal{F}$ and its inverse, and a characterization of the range were completed by Symes in [18]. Also see [2] for a perhaps cleaner proof of some of Symes's results. Symes also gave an inversion scheme and proved that it converged. The scheme was closely linked with the proof of all the other parts of the results but this scheme might not be the most efficient. Other efficient implementations of inversion schemes are given in [4] and it is understood that they are quite easy to implement.

Paper [3] is an excellent reference for a description and comparison of the various methods used for solving this problem. There are integral equation based schemes, Downward Continuation schemes, as in [18], which rely on the fact that in one space and time dimension the role of time and space may be interchanged, and a Layer Stripping scheme, as used in [19].

The references given above deal with a function $u(z, t)$ which satisfies a second-order PDE of the form

$$\eta \frac{\partial^2 u}{\partial t^2} - \frac{\partial}{\partial z} \left(\eta \frac{\partial u}{\partial z} \right) = 0 \tag{45}$$

instead of the first order system of PDEs. Here $\eta(z)$ is a positive function with $\eta(0) = 1$. The connection between the problem formulation in (42), (43) and the function u in the PDEs in the references given above is that L and R are the "left and right-going parts" of u. That is

$$L(z, t) = \sqrt{\eta(z)} \left(\frac{\partial u}{\partial t} + \frac{\partial u}{\partial z} \right) (z, t), \qquad R(z, t) = \sqrt{\eta(z)} \left(\frac{\partial u}{\partial t} - \frac{\partial u}{\partial z} \right) (z, t) \tag{46}$$

and

$$q(z) = \frac{\eta'(z)}{2\eta(z)}, \qquad \text{or equivalently,} \qquad \eta(z) = \exp \left(2 \int_0^z q(y) \, dy \right) \; .$$

For the piecewise-constant η case (one has to reinterpret what q means) this layer-stripping idea reduces to matrix manipulation and is explained quite well in [6].

2.7.2 Complex-valued q.

The complex-valued q case has been considered in many places, including [10], [17], and it arises in the study of solitons associated with the cubic, or nonlinear, Schrödinger equation – see [13]. However, the analysis in these articles is not as complete as for the real case. However, the techniques used by Symes in [18] carry over to the complex-valued case and similar results hold. Further work is being done on this and results will be sent to Greg Luther when it is complete.

3 Polarization Mode Dispersion Characterization: Determining the twist in an optical fiber

3.1 Introduction

A twisted birefringent fiber is probed by a wave at one end and the goal is to determine the twist in the fiber at each point along its length from the reflected wave. In [7] and [8], in the piecewise constant case, the forward and backward propagation is modeled as a multiplication by an orthogonal matrix at each interface. This suggests a model for the continuous case which was studied and analyzed in these articles. Greg Luther had wondered whether this model was accurate – particularly the reflection at an interface. A slightly different model is proposed below.

3.2 Basic formulation

Consider a fiber where z measures the distance from the left end of the fiber. Let $\mathbf{E}(z)$ be the electric field and $\mathbf{P}(z)$ the polarization at the point z units away from the left end of the fiber. We assume that $\mathbf{E}$ and $\mathbf{P}$ have no component along the fiber. $\mathbf{E}$ and $\mathbf{P}$ obey Maxwell's equations

$$\nabla^2\mathbf{E} - \nabla(\nabla\cdot\mathbf{E}) = \frac{1}{c^2}\frac{\partial^2\mathbf{E}}{\partial t^2} + \frac{1}{\epsilon_0 c^2}\frac{\partial^2\mathbf{P}}{\partial t^2}$$

where c is the speed of light in a vacuum and ϵ_0 is the permittivity of free space. Noting that $\mathbf{E}$ and $\mathbf{P}$ only depend on z and have no components in the $\mathbf{k}$ direction, Maxwell's equations reduce to

$$\mathbf{E}_{zz} = \frac{1}{c^2}\mathbf{E}_{tt} + \frac{1}{\epsilon_0 c^2}\mathbf{P}_{tt}\ . \tag{47}$$

At every point in the fiber, we can find two unit orthogonal vectors $\mathbf{v_1}(z)$ and $\mathbf{v_2}(z)$ perpendicular to the fiber, which are the principal directions of the fiber at that point. As the fiber twists along its length, the principal directions change. Since $\mathbf{v_1}$ and $\mathbf{v_2}$ are unit vectors in a plane and orthogonal, $d\mathbf{v_1}/dz$ is orthogonal to $\mathbf{v_1}(z)$ and hence parallel to $\mathbf{v_2}(z)$. Let $d\mathbf{v_1}/dz = \beta(z)\mathbf{v_2}$ for some real valued function $\beta(z)$. Also $d\mathbf{v_2}/dz$ is perpendicular to $\mathbf{v_2}$ so parallel to $\mathbf{v_1}$. But differentiating $\mathbf{v_1}\cdot\mathbf{v_2} = 0$ we have

$$0 = \frac{d\mathbf{v_1}}{dz}\cdot\mathbf{v_2} + \mathbf{v_1}\cdot\frac{d\mathbf{v_2}}{dz} = \beta + \mathbf{v_1}\cdot\frac{d\mathbf{v_2}}{dz}\ .$$

So $d\mathbf{v_2}/dz = -\beta\mathbf{v_1}$. Summarizing, the principal directions propagate through

$$\frac{d\mathbf{v_1}}{dz} = \beta\mathbf{v_2}, \qquad \frac{d\mathbf{v_2}}{dz} = -\beta\mathbf{v_1}\ . \tag{48}$$

Let E_1 and E_2 be the components of $\mathbf{E}$ along the principal directions, so $\mathbf{E} = E_1\mathbf{v_1} + E_2\mathbf{v_2}$. Then we assume that the polarization vector is related to the electric field via

$$\mathbf{P} = \epsilon_0(\alpha_1 E_1 \mathbf{v_1} + \alpha_2 E_2 \mathbf{v_2})$$

where α_1, α_2 are real constants. Substituting this in (47), (47) may be replaced by

$$\mathbf{E}_{zz} = \frac{1}{{c_1}^2} E_{1tt}\mathbf{v_1} + \frac{1}{{c_2}^2} E_{2tt}\mathbf{v_2} \ . \tag{49}$$

where we have assumed $1 + \alpha_i > 0$ and defined $c_i = 1/\sqrt{1+\alpha_i}$.

Now $\mathbf{E} = E_1\mathbf{v_1} + E_2\mathbf{v_2}$ combined with (48) gives

$$\begin{aligned}
\mathbf{E}_z &= E_{1z}\mathbf{v_1} + E_{2z}\mathbf{v_2} + \beta E_1\mathbf{v_2} - \beta E_2 \mathbf{v_1} = (E_{1z} - \beta E_2)\mathbf{v_1} + (E_{2z} + \beta E_1)\mathbf{v_2} \\
\mathbf{E}_{zz} &= (E_{1z} - \beta E_2)_z \mathbf{v_1} + (E_{2z} + \beta E_1)_z \mathbf{v_2} + \beta(E_{1z} - \beta E_2)\mathbf{v_2} - \beta(E_{2z} + \beta E_1)\mathbf{v_1} \\
&= \{(E_{1z} - \beta E_2)_z - \beta(E_{2z} + \beta E_1)\}\, \mathbf{v_1} \ + \ \{(E_{2z} + \beta E_1)_z + \beta(E_{1z} - \beta E_2)\}\, \mathbf{v_2} \ .
\end{aligned}$$

Using this in (49) and matching the components of $\mathbf{v_1}$ and $\mathbf{v_2}$ we have

$$(E_{1z} - \beta E_2)_z - \beta(E_{2z} + \beta E_1) = \frac{1}{{c_1}^2} E_{1tt} \tag{50}$$

$$(E_{2z} + \beta E_1)_z + \beta(E_{1z} - \beta E_2) = \frac{1}{{c_2}^2} E_{2tt} \ . \tag{51}$$

We will now convert this system of second order PDEs into a first order system in the left and right going waves. We start by defining the variables

$$M_1 = E_{1z} - \beta E_2, \qquad M_2 = E_{2z} + \beta E_1, \qquad M_3 = \frac{1}{c_1} E_{1t}, \qquad M_4 = \frac{1}{c_2} E_{2t} \ .$$

Then (50), (51) may be rewritten as

$$\begin{aligned}
c_1 M_{1z} - c_1 \beta M_2 &= M_{3t} \\
c_2 M_{2z} + c_2 \beta M_1 &= M_{4t}
\end{aligned}$$

combined with

$$\begin{aligned}
M_{1t} &= E_{1zt} - \beta E_{2t} = c_1 M_{3z} - \beta c_2 M_4 \\
M_{2t} &= E_{2zt} + \beta E_{1t} = c_2 M_{4z} + \beta c_1 M_3
\end{aligned}$$

So if $M = [M_1, M_2, M_3, M_4]^T$ then (50), (51) may be replaced by the first order system

$$M_t = AM_z + \beta BM \tag{52}$$

where

$$A = \begin{bmatrix} 0 & 0 & c_1 & 0 \\ 0 & 0 & 0 & c_2 \\ c_1 & 0 & 0 & 0 \\ 0 & c_2 & 0 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 0 & 0 & -c_2 \\ 0 & 0 & c_1 & 0 \\ 0 & -c_1 & 0 & 0 \\ c_2 & 0 & 0 & 0 \end{bmatrix}$$

Since A is symmetric, it may be diagonalized by an orthogonal matrix. In fact, $P^{-1}AP = D$ where

$$P = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \qquad D = \begin{bmatrix} c_1 & 0 & 0 & 0 \\ 0 & -c_1 & 0 & 0 \\ 0 & 0 & c_2 & 0 \\ 0 & 0 & 0 & -c_2 \end{bmatrix}.$$

Note that P is (almost) orthogonal in that $P^T P = 2I$. So if we introduce a new dependent variable $N = P^{-1}M$ then (52) may be rewritten as

$$N_t = DN_z + \beta CN \tag{53}$$

where

$$C = P^{-1}BP = \frac{1}{2}\begin{bmatrix} 0 & 0 & -c_1 - c_2 & -c_1 + c_2 \\ 0 & 0 & c_1 - c_2 & c_1 + c_2 \\ c_1 + c_2 & -c_1 + c_2 & 0 & 0 \\ c_1 - c_2 & -c_1 - c_2 & 0 & 0 \end{bmatrix}.$$

Let $N = [N_1, N_2, N_3, N_4]^T$ - if we backtrack we may obtain the relationship between the N_i and E_i. We obtain

$$2N_1 = E_{1z} - \beta E_2 + \frac{1}{c_1}E_{1t}, \qquad 2N_2 = E_{1z} - \beta E_2 - \frac{1}{c_1}E_{1t} \tag{54}$$

$$2N_3 = E_{2z} + \beta E_1 + \frac{1}{c_2}E_{2t}, \qquad 2N_4 = E_{2z} + \beta E_1 - \frac{1}{c_2}E_{2t} \tag{55}$$

and because of the form of D, N_1, N_2 are, loosely speaking, left and right going waves propagating with speed c_1, and N_3, N_4 left and right going waves propagating with speed c_2.

3.3 Goal

The fiber is probed by a right moving wave coming in from the left end and one of the components of the response is measured at the same end for a certain period of time - the goal is the recovery of the function $\beta(z)$.

We assume that the fiber is extended to the left of $z = 0$ and that $\beta(z) = 0$ for $z \leq 0$. The incoming wave is modeled as $E_1(t, z) = -H(c_1 t - z)$ and $E_2(t, z) = 0$ for $t < 0$ - note that the principal axes of the fiber do not change in the region $z \leq 0$. The measurement will consist of either $E_1(z = 0, t)$ or $E_2(z = 0, t)$ over a certain interval $t \in [0, T]$. The goal is to recover $\beta(z)$.

The initial conditions for $\mathbf{E}$ and the relationship between the E_i and N_i imply that

$$2N(t, z) = [0,\ 2\delta(c_1 t - z),\ -\beta H(c_1 t - z),\ -\beta H(c_1 t - z)]^T \qquad \text{for } t < 0\ .$$

Since $H(c_1 t - z)$ is supported in the region $c_1 t \geq z$, and $\beta(z) = 0$ for $z \leq 0$, we may conclude that

$$N(t, z) = [0,\ \delta(c_1 t - z),\ 0,\ 0]^T \qquad \text{for } t < 0\ . \tag{56}$$

Since $\beta(z) = 0$ for $z \leq 0$, (53) implies that

$$N_t = DN_z \qquad \text{for } z \leq 0,\ \text{all } t.$$

Since D is diagonal, on the region $z \leq 0$, N has the form

$$N = [n_1(c_1 t + z),\ n_2(c_1 t - z),\ n_3(c_2 t + z),\ n_4(c_2 t - z)]^T$$

for some functions n_i. The initial conditions (56), imply that $n_2(t) = \delta(t)$, $n_4(t) = 0$ and that $n_1(t)$ and $n_3(t)$ are supported in $t \geq 0$.

Hence $N(z,t)$ may be thought of as the solution of the initial boundary value problem

$$N_t = DN_z + \beta CN, \qquad z \geq 0,\ \text{all } t \tag{57}$$

$$N_2(z=0,t) = \delta(c_1 t), \qquad N_4(z=0,t) = 0, \qquad \text{all } t \tag{58}$$

$$N(z,t) = 0 \qquad \text{for } t < 0\ . \tag{59}$$

Our goal is that given $N_1(z=0,t)$ or $N_3(z=0,t)$ for t in some interval $[0,T]$, determine $\beta(z)$. Note that since $\beta(z) = 0$ for $z = 0$, (54) and (55) together with the boundary condition (58) imply that knowing N_1 or N_3 on $z = 0$ corresponds to knowing E_1 or E_2 on $z = 0$.

This problem seems similar to the grating design problem except that here there are two speeds of propagation and this makes the problem harder.

3.4 Literature

The phenomena of a twist in a birefringent fiber has been modeled and studied in various articles in journals on Electrical Engineering - see [7] and [8] for references. We have chosen to start with Maxwell's equation and the only approximation we have made is in the form of the polarization so the wave speeds are not frequency dependent.

Inverse problems for systems of hyperbolic equations in one space dimension, with multiple wave speeds have been studied in [1]. There they used techniques from Control Theory to determine several coefficients, however there were some gaps in the proof in that they claim that one of the relations is valid as a certain sequence of functions approaches the delta function but it is not clear that it is so. Our problem is a little simpler in that we attempt to recover only one function. A multidimensional linearized inverse problem about a known layered background with multiple wave speeds and the perturbation also being linear, was studied in [15].

3.5 Further thoughts on the model

The model for propogation and reflection of the waves might be developed along the following lines. In (47), suppose that the components in the (locally) principal directions satisfy $P_j = E_j/\epsilon_j$, $j = 1, 2$. Also take β to be a measure of the twist so that

$$\frac{d\mathbf{v_1}}{dz} = \beta \mathbf{v_2} \qquad \frac{d\mathbf{v_2}}{dz} = -\beta \mathbf{v_1}\ .$$

The components of the electrical field then satisfy equations of the form

$$\frac{\partial^2 E_1}{\partial z^2} - 2\beta \frac{\partial E_2}{\partial z} - \beta^2 E_1 - \beta' E_2 = (1-\alpha)\frac{\partial^2 E_1}{\partial t^2}\ ,$$

$$\frac{\partial^2 E_2}{\partial z^2} + 2\beta \frac{\partial E_1}{\partial z} - \beta^2 E_2 + \beta' E_1 = (1+\alpha)\frac{\partial^2 E_2}{\partial t^2}\ .$$

Here distance is scaled according to twist, $\beta = O(1)$ (and we expect $\beta' \equiv \frac{d\beta}{dz} = O(1)$), and time to make the basic wave speed equal to unity. If we look for a solution in the form of a wave packet,

$$\begin{pmatrix} E_1 \\ E_2 \end{pmatrix} = \begin{pmatrix} A_1(z,t) \\ A_2(z,t) \end{pmatrix} e^{ik(z-t)}, \tag{60}$$

and take a high wave number, $k \gg 1$, we find, approximately,

$$\frac{\partial^2 A_1}{\partial z^2} + 2ik\frac{\partial A_1}{\partial z} - 2\beta\frac{\partial A_2}{\partial z} - 2ik\beta A_2 - \beta^2 A_1 - \beta' A_2 = \frac{\partial^2 A_1}{\partial t^2} + 2i\alpha k\frac{\partial A_1}{\partial t} + \alpha k^2 A_1 - 2ik\frac{\partial A_1}{\partial t},$$

$$\frac{\partial^2 A_2}{\partial z^2} + 2ik\frac{\partial A_2}{\partial z} + 2\beta\frac{\partial A_1}{\partial z} + 2ik\beta A_1 - \beta^2 A_2 + \beta' A_1 = \frac{\partial^2 A_2}{\partial t^2} - 2i\alpha k\frac{\partial A_2}{\partial t} - \alpha k^2 A_2 - 2ik\frac{\partial A_2}{\partial t}.$$

A change to travelling-wave co-ordinates, writing $\zeta = z - t$, and an assumption that ϵ_1 and ϵ_2 are close, so that $\gamma \equiv \alpha k/2$ is $O(1)$, gives

$$\begin{aligned}
\frac{\partial A_1}{\partial t} \sim\ & -i\gamma A_1 + \beta A_2 \\
& + \frac{1}{k}\left(-i\beta\frac{\partial A_2}{\partial \zeta} - \frac{i\beta^2}{2}A_1 - \frac{i\beta'}{2}A_2 - \frac{i}{2}\frac{\partial^2 A_1}{\partial t^2} + i\frac{\partial^2 A_1}{\partial t\partial\zeta} + 2\gamma\frac{\partial A_1}{\partial t} - 2\gamma\frac{\partial A_1}{\partial \zeta}\right), \\
\frac{\partial A_2}{\partial t} \sim\ & -\beta A_1 + i\gamma A_2 \\
& + \frac{1}{k}\left(i\beta\frac{\partial A_1}{\partial \zeta} - \frac{i\beta^2}{2}A_2 + \frac{i\beta'}{2}A_1 - \frac{i}{2}\frac{\partial^2 A_2}{\partial t^2} + i\frac{\partial^2 A_2}{\partial t\partial\zeta} - 2\gamma\frac{\partial A_2}{\partial t} + 2\gamma\frac{\partial A_2}{\partial \zeta}\right).
\end{aligned} \tag{61}$$

To leading order,

$$\begin{pmatrix} A_1 \\ A_2 \end{pmatrix} = \boldsymbol{U}_F(t)\mathbf{a}(\zeta) = \boldsymbol{U}_F(t)\mathbf{a}(z-t), \tag{62}$$

where $\mathbf{a}(\zeta)$ is determined by the initial pulse of light supplied at $z = 0$ around $t = 0$, and the unitary matrix $\boldsymbol{U}_F$ is given by

$$\frac{d\boldsymbol{U}_F}{dz} = \boldsymbol{M}_F\boldsymbol{U}_F, \qquad \boldsymbol{M}_F = \begin{pmatrix} -i\gamma & \beta \\ -\beta & i\gamma \end{pmatrix}, \qquad U_F(0) = I$$

(*c.f.* (27) and (33) in §2).

Now (62) will not be an exact solution for the coupled wave equations; there must also be a backwards wave,

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} e^{-ik(z+t)}.$$

Here, the magnitude of this backward-travelling, or "left-going", wave is expected to be exponentially small in k; during the workshop, such a reflected wave train was taken to be produced by some other means – and in a manner independent of position z (so β *etc.* had no rôle in actual reflection events).

Ahead of the wave packet, *i.e.* in $z > t$, $B_1 = B_2 = 0$, while behind it, in $z < t$, to leading order,

$$\left(\frac{\partial}{\partial t} - \frac{\partial}{\partial z}\right)\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \boldsymbol{M}_B\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \qquad \boldsymbol{M}_B = \begin{pmatrix} -i\gamma & -\beta \\ \beta & i\gamma \end{pmatrix} = \boldsymbol{M}_F^T.$$

Integrating along a characteristic $z = 2Z - t$, so

$$\frac{d}{dt}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = -\frac{d}{dz}\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \boldsymbol{M}_B \begin{pmatrix} B_1 \\ B_2 \end{pmatrix},$$

gives the value at $z = 0$ in terms of another unitary matrix $U_B(Z)$:

$$\left.\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}\right|_{z=0} = \boldsymbol{U}_B(Z) \left.\begin{pmatrix} B_1 \\ B_2 \end{pmatrix}\right|_{z=Z},$$

with

$$\boldsymbol{U}_B(Z) = \boldsymbol{R}_Z(0)\,, \qquad \frac{d\boldsymbol{R}_Z}{dz} = -\boldsymbol{M}_B\boldsymbol{R}_Z\,, \qquad \boldsymbol{R}_Z(Z) = \boldsymbol{I}\,.$$

This, in principle, determines the electric field at the end of the fibre, $z = 0$, at time $t = 2Z$, as long as the field on reflection at $z = Z$, $(B_1, B_2)|_{z=Z}$, is known – or can be found (simply) from the original propagating wave packet, *i.e.* from (A_1, A_2).

The reflection should, ideally, be obtained by solving the wave equations in the vicinity of the forwards, or "right-going", wave. This might be expected to lead to a reflection at $z = Z$ which depends upon the local values of α, β and β', rather than having an assumed (small) constant of proportionality between (A_1, A_2) and (B_1, B_2). This would make the interval, or "layer-stripping", method of finding β *etc.* at successively larger values of $z = Z_n$ from measurements of $(B_1(0, 2Z_n), B_2(0, 2Z_n))$ considerably more difficult.

($\boldsymbol{U}_B(Z)$ can also be found slightly differently from above. If the unitary matrix $\boldsymbol{R}_B$ is given by

$$\frac{d\boldsymbol{R}_B}{dz} = -\boldsymbol{M}_B\boldsymbol{R}_B\,, \qquad \boldsymbol{R}_B(0) = \boldsymbol{I}\,,$$

then

$$\boldsymbol{U}_B(Z) = \boldsymbol{R}_B(Z)^{-1} = \boldsymbol{R}_B^*(Z) \equiv \overline{\boldsymbol{R}_B(Z)^T}\,.$$

Of course, since $\boldsymbol{U}_B$ is unitary, $\boldsymbol{U}_B^{-1} = \overline{\boldsymbol{U}_B^T} = \boldsymbol{R}_B(Z)$ satisfies

$$\frac{d}{dz}\overline{\boldsymbol{U}_B^T} = -\boldsymbol{M}_B\overline{\boldsymbol{U}_B^T} \quad \text{and} \qquad \frac{d}{dz}\boldsymbol{U}_B = -\boldsymbol{U}_B\overline{\boldsymbol{M}_B^T} = \boldsymbol{U}_B\boldsymbol{M}_F\,.$$

The paper [7] gives an alternative description for these transition matrices as a rotation in three dimensions.)

References

[1] M. Belishev, A. Blagovestchenskii & S. Ivanov, *The two velocity dynamical system: boundary control of waves and inverse problems*, Wave Motion, **25**, 1997, 83-107.

[2] B.L. Browning, Time and Frequency Domain Scattering for the One Dimensional Wave Equation, PhD Thesis, University of Washington, Seattle, 1999.

[3] A.M. Bruckstein, B.C. Levy & T. Kailath, *Differential methods in inverse scattering*, SIAM Jl. Appl. Maths., **45**(2), 1985, 312-335.

[4] K. Bube & R. Burridge, *The one-dimensional inverse problem of reflection seismology*, SIAM Rev., **25**, 1983, 497-559.

[5] C.R. Chester, Techniques in Partial Differential Equations, McGraw-Hill, New York, 1971.

[6] J. Claerbout, Fundamentals of Geophysical Data Processing, Blackwell Scientific Publications, Oxford, 1985.

[7] F. Corsi, A. Galtarossa & L. Palmieri, *Polarization mode dispersion characterization of single-mode optical fiber using backscattering technique*, Jl. of Lightwave Technology, **16**(10), 1998, 1832-1843.

[8] F. Corsi, A. Galtarossa & L. Palmieri, *Analytical Treatment of Polarization Dispersion In single mode fibers by means Of backscattering signal*, preprint.

[9] R. Courant & D. Hilbert, Methods of Mathematical Physics Vol.I, Wiley, New York, 1953.

[10] R. Feced, M.N. Zervas & M.A. Muriel, *An efficient inverse scattering algorithm for the design of nonuniform fiber Bragg gratings*, IEEE Jl. of Quantum Electronics, **35**(8), 1999, 1105-1115.

[11] S.L. He, S.I. Kabanikhin, V.G. Romanov & S. Strom, *Analysis of the Green's function approach to one dimensional inverse problems. I - one parameter reconstruction.* Jl. Math. Phys., **34**(12), 1993, 5724-5746.

[12] S.L. He, V.G. Romanov & S. Strom, *Analysis of the Green's function approach to one dimensional inverse problems. II - simulateneous reconstruction of two parameters.* Jl. Math. Phys., **35**(5), 1994, 2315-2335.

[13] G.L. Lamb, Jr., Elements of Soliton Theory, Wiley, New York, 1980.

[14] A.C. Pipkin, A Course on Integral Equations, Springer, New York, 1991.

[15] P. Sacks & W. Symes *Recovery of the elastic parameters of a layered half space*, Geophys. J. R. Astr. Soc., **88**, 1987, 593-620.

[16] J.E. Sipe, L. Poladian & C. Martijn de Sterke, *Propagation through nonuniform grating structures*, Jl. Opt. Soc. Am. A, **11**(4), 1994.

[17] G.-H. Song & S.-Y. Shin, *Design of corrugated waveguide filters by the Gelfand-Levitan-Marchento inverse-scattering method*, J. Optical Society of America A, 2(11), 1985, 1905-1915.

[18] W.W. Symes, *Impedance profile inversion via the first transport equation*, Jl. Math. Anal., **94**(2), 1983, 435-453.

[19] J. Sylvester, D.P. Winebrenner & F. Gylys-Colwell, *Layer stripping for the Helmholtz equation*, SIAM Jl. Appl. Math., **56**, 1996, 736-754.

[20] E.T. Whittaker & G.N. Watson, A Course of Modern Analysis, C.U.P., 1935.