INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/842-5

Conference on
Ultrafast Transmission Systems in Optical Fibres

13 - 17 February 1995

Solitons in Randomly Birefringent Fibers

J.P. Gordon

AT&T Bell Laboratories
Holmdel, New Jersey
U.S.A.

Propagation in birefringent fibers (using the Pauli vector)

James P. Gordon

In dealing with problems of propagation in birefringent fibers, such as polarization dispersion, I have found it helpful to use a combination of vector and matrix notation in what is perhaps a novel way.

A field with two polarizations is represented by a column matrix, and its conjugate transpose by a row matrix; thus

$$\tilde{U}(z,\omega) = \begin{bmatrix} \tilde{u}_x(z,\omega) \\ \tilde{u}_y(z,\omega) \end{bmatrix} \quad ; \quad \tilde{U}^\dagger(z,\omega) = \begin{bmatrix} \tilde{u}_x^*(z,\omega) & \tilde{u}_y^*(z,\omega) \end{bmatrix} \tag{1.1}$$

The tilde over $\tilde{U}$ and $\tilde{u}$ indicates that they are monochromatic fields.

The Stokes vector corresponding to such a field exists in a three dimensional space, which we shall call Stokes space. It is represented by $\vec{S}$ and has the components

$$\begin{aligned} S_1 &= |\tilde{u}_x|^2 - |\tilde{u}_y|^2 \\ S_2 &= \tilde{u}_x^*\tilde{u}_y + \tilde{u}_y^*\tilde{u}_x \\ S_3 &= -i\,\tilde{u}_x^*\tilde{u}_y + i\,\tilde{u}_y^*\tilde{u}_x \end{aligned} \tag{1.2}$$

The three unit vectors along the three axes of Stokes space we label $\hat{e}_1$, $\hat{e}_2$ and $\hat{e}_3$. From Eqs.(1.2) one can show that the length of the Stokes vector is given by $S_0 = \tilde{U}^\dagger\tilde{U} = |\tilde{u}_x|^2 + |\tilde{u}_y|^2$. Linearly polarized fields, for which $\tilde{u}_y/\tilde{u}_x$ is real, have their Stokes vectors in the $(\hat{e}_1,\hat{e}_2)$ plane, while circularly polarized fields have their Stokes vectors along the $\hat{e}_3$ axis.

The three relations in Eqs.(1.2) can be more succinctly expressed using the three Pauli matrices, familiar to physicists from the quantum mechanics of two state systems. Reordered to fit the present context, these are

$$\boldsymbol{\sigma}_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad ; \quad \boldsymbol{\sigma}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad ; \quad \boldsymbol{\sigma}_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{1.3}$$

With the Pauli matrices, Eqs.(1.2) can be reexpressed as

$$S_j = \tilde{U}^\dagger\boldsymbol{\sigma}_j\tilde{U} \qquad j = 1,2,3 \tag{1.4a}$$

Eq.(1.4a) suggests the notation we advance here, and that is to treat the three Pauli matrices as the three components of a vector in Stokes space. Thus, we define the "matrix-vector"

$$\vec{\boldsymbol{\sigma}} \equiv \boldsymbol{\sigma}_1\hat{e}_1 + \boldsymbol{\sigma}_2\hat{e}_2 + \boldsymbol{\sigma}_3\hat{e}_3$$

With this device, one can reexpress Eq(1.4a) vectorially as

$$\vec{S} = \tilde{U}^\dagger\vec{\boldsymbol{\sigma}}\tilde{U} \tag{1.4b}$$

The properties of the Pauli matrices are important for what follows. They, along with the unit matrix **I**, form a complete set of Hermitian 2 by 2 matrices. A Hermitian matrix is one which is equal to its complex conjugate transpose, that is, its Hermitian conjugate. It has real eigenvalues. Any 2 by 2 Hermitian matrix may be expanded in the form

$$a_0\mathbf{I} + a_1\boldsymbol{\sigma}_1 + a_2\boldsymbol{\sigma}_2 + a_3\boldsymbol{\sigma}_3 = a_0\mathbf{I} + \vec{a}\cdot\vec{\boldsymbol{\sigma}} \tag{1.5a}$$

where the coefficients a_j are real, and the last form, using the vector dot product, takes advantage of the matrix-vector notation , i.e.

$$\vec{a}\cdot\vec{\boldsymbol{\sigma}} \equiv a_1\boldsymbol{\sigma}_1 + a_2\boldsymbol{\sigma}_2 + a_3\boldsymbol{\sigma}_3 = \begin{bmatrix} a_1 & a_2 - ia_3 \\ a_2 + ia_3 & -a_1 \end{bmatrix} \tag{1.5b}$$

In context, the vector $\vec{a}$ in Eqs.(1.5) is a vector in Stokes space. All such vectors are assumed to have real components unless the contrary is explicitly specified. The components of $\vec{\boldsymbol{\sigma}}$ are Hermitian, which is the matrix equivalent of real.

The Pauli matrices obey the well known multiplication rules,

$$\sigma_j^2 = \mathbf{I} \quad ; \quad \sigma_j\sigma_k = -\sigma_k\sigma_j = i\sigma_l \tag{1.6}$$

where (j,k,l) can be any cyclic permutation of (1,2,3). Using Eqs.(1.6), any function of the $\boldsymbol{\sigma}$ matrices that can be expanded in a power series can be reduced to an expression linear in the $\boldsymbol{\sigma}$ matrices and the unit matrix $\mathbf{I}$. Examples which will be used below, expressed in vector notation, include

$$\vec{\sigma}(\vec{a}\cdot\vec{\sigma}) = \vec{a}\,\mathbf{I} + i\vec{a}\times\vec{\sigma} \tag{1.7a}$$

$$(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma}) = (\vec{a}\cdot\vec{b})\,\mathbf{I} + i(\vec{a}\times\vec{b})\cdot\vec{\sigma} \tag{1.7b}$$

$$(\vec{a}\cdot\vec{\sigma})\vec{\sigma}(\vec{a}\cdot\vec{\sigma}) = 2\vec{a}(\vec{a}\cdot\vec{\sigma}) - a^2\vec{\sigma} \tag{1.7c}$$

$$\exp(i\vec{a}\cdot\vec{\sigma}) = \cos(a)\,\mathbf{I} + i\sin(a)\,\hat{a}\cdot\vec{\sigma} \tag{1.7d}$$

In Eqs.(1.7), the vectors $\vec{a}$ and $\vec{b}$ can be any vectors in Stokes space. Examples of both dot and cross vector products occur. In Eq.(1.7a), for example, the $\hat{e}_1$ component reads $\sigma_1(\vec{a}\cdot\vec{\sigma}) = a_1\mathbf{I} + i(a_2\sigma_3 - a_3\sigma_2)$, which follows from Eqs.(1.5b) and (1.6). In Eqs.(1.7c) and (1.7d), a is the length of the vector $\vec{a}$, while $\hat{a}$ is the unit vector $\vec{a}/a$. This notation will also be applied to other vectors. The final example (1.7d) derives from the result of Eq.(1.7b) when $\vec{b}=\vec{a}$, namely $(\vec{a}\cdot\vec{\sigma})^2 = a^2\mathbf{I}$. The equations conjugate to Eqs.(1.7) also apply. The equation conjugate to Eq.(1.7a) is $(\vec{a}\cdot\vec{\sigma})\vec{\sigma} = \vec{a}\,\mathbf{I} - i\vec{a}\times\vec{\sigma}$. Interchanging $\vec{a}$ and $\vec{b}$ conjugates Eq.(1.7b), while Eq.(1.7c) is its own conjugate, and replacing i by $-i$ conjugates Eq.(1.7d). Note that $\vec{a}\cdot\vec{\sigma}$ does not commute with $\vec{b}\cdot\vec{\sigma}$ unless $\vec{a}\times\vec{b}=0$.

Eigenvalues and eigenstates

With any Hermitian matrix of the form $\vec{a}\cdot\vec{\sigma}$ one can associate a pair of orthogonal eigenstates of the field and a new basis set of orthogonal unit vectors in Stokes space, one of which is the unit vector $\hat{a}$. These facilitate the discussion of problems in regard to birefringent propagation. Since $(\vec{a}\cdot\vec{\sigma})^2 = a^2\mathbf{I}$, it follows that the eigenvalues of $\vec{a}\cdot\vec{\sigma}$ are a and $-a$. With each eigenvalue is associated an eigenstate. Let $\tilde{U}_+$ and $\tilde{U}_-$ be the two normalized eigenstates of $(\vec{a}\cdot\vec{\sigma})$. They have Stokes vectors of unit length, and satisfy the

equations

$$\vec{a}\cdot\vec{\sigma}\,\tilde{U}_+ = a\tilde{U}_+ \quad ; \quad \vec{a}\cdot\vec{\sigma}\,\tilde{U}_- = -a\tilde{U}_- \tag{2.1}$$

Using Eq.(1.1), one can show from Eqs.(2.1) that the eigenstate $\tilde{U}_+$ has $\tilde{u}_y/\tilde{u}_x = (a_2+ia_3)/(a+a_1)$, while the eigenstate $\tilde{U}_-$ has $\tilde{u}_y/\tilde{u}_x = -(a+a_1)/(a_2-ia_3)$, and that $\tilde{U}^\dagger_+\tilde{U}_- = 0$, so that they are orthogonal to one another. (Eigenstates with different eigenvalues are always orthogonal to one another.) The eigenstates are linearly polarized if $a_3=0$, circularly polarized if $a_1=a_2=0$, and otherwise elliptically polarized. The eigenstates have independent phases, which can be chosen arbitrarily. Any field can be expanded in terms of them, so they are complete. These properties are represented by the equations

$$\tilde{U}_+^\dagger\tilde{U}_+ = \tilde{U}_-^\dagger\tilde{U}_- = 1 \qquad \text{(normalization)} \tag{2.2a}$$

$$\tilde{U}_+^\dagger\tilde{U}_- = \tilde{U}_-^\dagger\tilde{U}_+ = 0 \qquad \text{(orthogonality)} \tag{2.2b}$$

$$\tilde{U}_+\tilde{U}_+^\dagger + \tilde{U}_-\tilde{U}_-^\dagger = \mathbf{I} \qquad \text{(completeness)} \tag{2.2c}$$

To generate the new basis set of vectors in Stokes space, we use Eq.(1.7a) and any two fields $\tilde{U}_\alpha$ and $\tilde{U}_\beta$ to obtain

$$\tilde{U}_\alpha^\dagger\vec{\sigma}(\vec{a}\cdot\vec{\sigma})\tilde{U}_\beta = \vec{a}(\tilde{U}_\alpha^\dagger\tilde{U}_\beta) + i\vec{a}\times(\tilde{U}_\alpha^\dagger\vec{\sigma}\tilde{U}_\beta)$$

Now replacing $\tilde{U}_\alpha$ and $\tilde{U}_\beta$ by the eigenstates in various ways, and using Eqs.(2.1) and (2.2), we find the results

$$(\tilde{U}_+^\dagger\vec{\sigma}\tilde{U}_+) = \vec{a}/a \tag{2.3a}$$

$$(\tilde{U}_-^\dagger\vec{\sigma}\tilde{U}_-) = -\vec{a}/a \tag{2.3b}$$

$$(\tilde{U}_-^\dagger\vec{\sigma}\tilde{U}_+) = i(\vec{a}/a)\times(\tilde{U}_-^\dagger\vec{\sigma}\tilde{U}_+) \tag{2.3c}$$

along with the conjugate of Eq.(2.3c). Eqs. (2.3a) and (2.3b) demonstrate that the Stokes vectors of the normalized field eigenstates are the unit vectors $\hat{a}$ and $-\hat{a}$. (The Stokes vectors of orthogonal fields are always antiparallel.) From Eq.(2.3c) it follows that the complex vector $\tilde{U}_-^\dagger\vec{\sigma}\tilde{U}_+$ is perpendicular to $\hat{a}$. If we make the definitions

$$\hat{a}_1 \equiv \hat{a} = \tilde{U}_+^\dagger \vec{\sigma} \tilde{U}_+ \tag{2.4a}$$

$$\hat{a}_2 + i\hat{a}_3 \equiv \tilde{U}_-^\dagger \vec{\sigma} \tilde{U}_+ \tag{2.4b}$$

then the real and imaginary parts of Eq.(2.3c) yield $\hat{a}_2 = -\hat{a}_1 \times \hat{a}_3$, and $\hat{a}_3 = \hat{a}_1 \times \hat{a}_2$. One can show using the completeness relation Eq.(2.2c) that the vectors $\hat{a}_2$ and $\hat{a}_3$ are also unit vectors, so the three unit vectors $\hat{a}_j$ form a new right handed basis set for the Stokes space. (Note that the three unit vectors $\hat{a}_j$ are quite different from the three components of $\vec{a}$ which appeared above.) If one expands an arbitrary field state in terms of the eigenstates of $\vec{a}\cdot\vec{\sigma}$

$$\tilde{U} = f\tilde{U}_+ + g\tilde{U}_- \tag{2.5a}$$

then one finds the Stokes vector $\vec{S} = \tilde{U}^\dagger \vec{\sigma} \tilde{U}$ expanded into the form

$$\vec{S} = (|f|^2 - |g|^2)\hat{a}_1 + (f^* g + f g^*)\hat{a}_2 + (-if^* g + ifg^*)\hat{a}_3 \tag{2.5b}$$

which is evidently a representation of $\vec{S}$ in a rotated coordinate system based on $\hat{a}_1$. The original coordinate system is recreated using $\hat{a}_1 = \hat{e}_1$, $\tilde{U}_+ = \begin{pmatrix}1\\0\end{pmatrix}$, and $\tilde{U}_- = \begin{pmatrix}0\\1\end{pmatrix}$. A phase change of $\tilde{U}_-$ with respect to $\tilde{U}_+$ rotates $\hat{a}_2$ and $\hat{a}_3$ around the axis of $\hat{a}_1$.

In the following sections, we shall have repeated occasions to use the results of this exercise.

Birefringent propagation

Fibers of good quality for optical transmission always have some random residual birefringence. The part of the propagation equation which describes the local wavenumber and its birefringence can be written in vector form as

$$\frac{\partial \tilde{U}}{\partial z} = i(\beta_0 \mathbf{I} + \frac{1}{2}\vec{\beta}\cdot\vec{\sigma})\,\tilde{U} \tag{3.1}$$

Here β_0 is the mean wavenumber, while the birefringence vector $\vec{\beta}$ is a vector in Stokes space. The Hermitian form of the operator in parentheses on the right is required to

preserve power, and therefore the length $\tilde{U}^\dagger \tilde{U}$ of the Stokes vector. The magnitude of the birefringence is the difference of the two eigenvalues of $\vec{\beta}\cdot\vec{\sigma}/2$, and is therefore the length β of the the birefringence vector $\vec{\beta}$. As we have also seen in the last section, the Stokes vectors of the two orthogonal field eigenstates of the birefringence are the unit vectors $\hat{\beta}$ and $-\hat{\beta}$.

Corresponding to Eq.(3.1) is the equation for the propagation of the Stokes vector. Differentiating $\vec{S}$ in Eq.(1.4b) with respect to z and using Eq.(3.1) and its conjugate, one gets

$$\frac{\partial \vec{S}}{\partial z} = \tilde{U}^\dagger \left[\frac{-i}{2}(\vec{\beta}\cdot\vec{\sigma})\vec{\sigma} + \frac{i}{2}\vec{\sigma}(\vec{\beta}\cdot\vec{\sigma}) \right] \tilde{U} \tag{3.2}$$

With the aid of Eq.(1.7a) and its conjugate, this becomes

$$\frac{\partial \vec{S}}{\partial z} = -\tilde{U}^\dagger \vec{\beta}\times\vec{\sigma}\tilde{U} = -\vec{\beta}\times\vec{S} \tag{3.3}$$

Thus, as is well known, the propagation of the Stokes vector consists of precession around the birefringence vector. The propagation preserves angles between the Stokes vectors corresponding to differently polarized fields. The result $\partial(\vec{S}_a\cdot\vec{S}_b)/\partial z = 0$ follows from Eq.(3.3). Generalizing the second equality of Eq.(3.3), we note that if $\overleftrightarrow{R}\cdot\vec{\sigma}$ represents any linear operation on the vector $\vec{\sigma}$, such as a rotation, then consideration of its components $(\overleftrightarrow{R}\cdot\vec{\sigma})_j = \sum_k R_{jk}\sigma_k$ shows, with Eq.(1.4a), that

$$\tilde{U}^\dagger \overleftrightarrow{R}\cdot\vec{\sigma}\tilde{U} = \overleftrightarrow{R}\cdot\vec{S} \tag{3.4}$$

The vector cross product of Eq.(3.3) is one such linear operation.

Propagation over a finite distance

Linear propagation over a finite distance in a birefringent fiber, from z_1 to z, say, is generally described by a 2 by 2 Jones matrix, which we can label $\mathbf{T}(z, z_1)$. Thus we have

$$\tilde{U}(z) = \mathbf{T}(z,z_1)\tilde{U}(z_1) \tag{4.1}$$

For loss free propagation, any Jones matrix can be expressed in the general unitary form

$$\mathbf{T}(z,z_1) = \exp\left[i(\theta_0\mathbf{I} + \frac{1}{2}\vec{\theta}\cdot\vec{\sigma})\right] = \exp(i\theta_0)\exp(\frac{i}{2}\vec{\theta}\cdot\vec{\sigma}) \tag{4.2a}$$

By expanding $\exp(i\vec{\theta}\cdot\vec{\sigma}/2)$ using Eq.(1.7d), one may verify that the Jones matrix can also be written in the form

$$\mathrm{T}(z,z_1) = \exp(i\theta_0)\begin{bmatrix} t & r \\ -r^* & t^* \end{bmatrix} \tag{4.2b}$$

where $|t|^2+|r|^2=1$. In this form the matrix elements are the Cayley-Klein parameters of mechanics, used there also as a representation of a rigid rotation. A unitary matrix is one whose inverse is equal to its Hermitian conjugate, as is evident from the form of Eq.(4.2a). This property is necessary to preserve the value of $\tilde{U}^\dagger\tilde{U}$. As we shall see, such a Jones matrix corresponds to a finite rotation in Stokes space, of amount $-\theta$ around the direction of the unit vector $\hat{\theta}$, where as usual $\vec{\theta}=\theta\hat{\theta}$.

If $\hat{\beta}=\vec{\beta}/\beta$ is independent of z in Eq.(3.1), then that equation can be integrated to give $\mathbf{T}(z,z_1)$ with

$$\theta_0 = \int_{z_1}^{z} dz\beta_0 \quad ; \quad \vec{\theta} = \int_{z_1}^{z} dz\beta\,\hat{\beta} \tag{4.3}$$

This case corresponds to a finite rotation around the direction $\hat{\beta}$. In the more general case, where $\hat{\beta}$ varies with position along the fiber, the Jones matrix will still have the form of Eq.(4.2a), but in the this case the direction of $\vec{\theta}$ will be frequency dependent. For example, concatenated sections of fiber, each with its own constant value of $\hat{\beta}$, will have a total Jones matrix which is the ordered product of the Jones matrices for each section. The net rotation is the ordered product of the rotations corresponding to each section, and will depend in direction and magnitude on both the directions and magnitudes of the rotations due to each section and on the order in which they occur.

It is useful to define $\mathbf{T}(z_1,z)$ as the inverse of $\mathbf{T}(z,z_1)$, for then the concatenation rule

$$\mathbf{T}(z,z_1) = \mathbf{T}(z,z_2)\,\mathbf{T}(z_2,z_1) \tag{4.4}$$

holds whether or not z_2 is between z and z_1.

The corresponding propagation of the Stokes vector is described by a finite rotation, a real 3×3 Mueller matrix, which we label $\overleftrightarrow{M}$ in dyadic notation to distinguish it from a Jones matrix. Thus we have the relation

$$\vec{S}(z) = \overleftrightarrow{M}(z,z_1)\cdot\vec{S}(z_1) \tag{4.5}$$

A general relation between the Jones and Mueller matrices can be established. Eq.(4.5) expands , using Eqs.(1.4b), (3.4) and (4.1), to read

$$\tilde{U}^\dagger(z_1)\mathbf{T}(z_1,z)\vec{\sigma}\mathbf{T}(z,z_1)\tilde{U}(z_1) = \tilde{U}^\dagger(z_1)\overleftrightarrow{M}(z,z_1)\cdot\vec{\sigma}\tilde{U}(z_1)$$

Since the field state $\tilde{U}(z_1)$ is arbitrary, one can extract the included equation

$$\mathbf{T}(z_1,z)\vec{\sigma}\,\mathbf{T}(z,z_1) = \overleftrightarrow{M}(z,z_1)\cdot\vec{\sigma} \tag{4.6}$$

In this equation, the Jones matrices $\mathbf{T}$ operate on the individual Pauli matrices σ_j, while the Mueller matrix rotates $\vec{\sigma}$ as a vector in Stokes space. Using Eq.(4.2a) for $\mathbf{T}(z,z_1)$, its inverse for $\mathbf{T}(z_1,z)$, and Eqs.(7), the left side of Eq.(4.6) expands to give

$$\overleftrightarrow{M}(z,z_1)\cdot\vec{\sigma} = \vec{\sigma} - \sin\theta(\hat{\theta}\times\vec{\sigma}) + (1-\cos\theta)\hat{\theta}\times(\hat{\theta}\times\vec{\sigma}) \tag{4.7}$$

Thus $\overleftrightarrow{M}(z,z_1)$ here represents a precession of $\vec{\sigma}$, considered as an arbitrary vector in Stokes space, through an angle of $-\theta$ around the direction of $\hat{\theta}$. Indeed, by virtue of Eq.(3.4), Eq.(4.7) holds when $\vec{\sigma}$ is replaced by any Stokes vector. If $\overleftrightarrow{M}(z_1,z)$ is defined as the inverse (also the transpose) of $\overleftrightarrow{M}(z,z_1)$, then Eqs.(4.5) and (4.6) hold for all values of z and z_1, and the Mueller matrices have the same concatination rule as the Jones matrices. In Eqs.(4.2a) and (4.7), the inverse matrices are obtained by substituting $-\vec{\theta}$ for $\vec{\theta}$ and $-\theta_0$ for θ_0.

A useful corollary to Eq.(4.6) is

$$\mathbf{T}(z_1,z)\vec{a}\cdot\vec{\sigma}\,\mathbf{T}(z,z_1) = \vec{a}\cdot(\overleftrightarrow{M}(z,z_1)\cdot\vec{\sigma}) = (\overleftrightarrow{M}(z_1,z)\cdot\vec{a})\cdot\vec{\sigma} \tag{4.8}$$

where $\vec{a}$ can be any vector in Stokes space, and the last step uses the invariance of a vector dot product to a common rotation of both vectors.

The properties of the Jones and Mueller matrices can also be examined using the eigenstates of $\vec{\theta}\cdot\vec{\sigma}$. Since the eigenvalues of $\vec{\theta}\cdot\vec{\sigma}$ are θ and $-\theta$, the eigenvalues of the Jones matrix $\mathbf{T}(z,z_1)$ are $\exp(i\theta_0+i\theta/2)$ and $\exp(i\theta_0-i\theta/2)$. As before, we label the eigenstates of $\vec{\theta}\cdot\vec{\sigma}$ here $\tilde{U}_+$ and $\tilde{U}_-$, and generate a basis set for the Stokes space by

$$\hat{\theta}_1 = \hat{\theta} = \tilde{U}_+{}^\dagger\vec{\sigma}\tilde{U}_+ \quad ; \quad \hat{\theta}_2 + i\hat{\theta}_3 = \tilde{U}_-{}^\dagger\vec{\sigma}\tilde{U}_+ \tag{4.9}$$

The expansion of a field at location z_1 in these eigenstates in the form

$$\tilde{U}(z_1) = a\tilde{U}_+ + b\tilde{U}_- \tag{4.10a}$$

yields, as in Eq.(2.5b), the Stokes vector in the form

$$\vec{S}(z_1) = (|a|^2-|b|^2)\hat{\theta}_1 + a^*b(\hat{\theta}_2-i\hat{\theta}_3) + ab^*(\hat{\theta}_2+i\hat{\theta}_3) \tag{4.10b}$$

After propagating to z, the field transforms to

$$\tilde{U}(z) = \exp(i\theta_0)(a\exp(i\theta/2)\tilde{U}_+ + b\exp(-i\theta/2)\tilde{U}_-) \tag{4.11a}$$

and so the Stokes vector transforms to

$$\begin{aligned}\vec{S}(z) = (|a|^2-|b|^2)\hat{\theta}_1 &+ \exp(-i\theta)a^*b(\hat{\theta}_2-i\hat{\theta}_3)\\ &+ \exp(i\theta)ab^*(\hat{\theta}_2+i\hat{\theta}_3)\end{aligned} \tag{4.11b}$$

One can see from the transformation of $\vec{S}$ that the Mueller matrix has one real eigenvector, $\hat{\theta}$, with eigenvalue 1, and two complex eigenvectors, $(\hat{\theta}_2-i\hat{\theta}_3)/\sqrt{2}$, and $(\hat{\theta}_2+i\hat{\theta}_3)/\sqrt{2}$, with eigenvalues $\exp(-i\theta)$, and $\exp(i\theta)$, respectively. These results are consistent with Eq.(4.7) for the Mueller matrix, as one can see by substituting each eigenvector for $\vec{\sigma}$ in that equation.

The eigenstates of the Jones matrix are those two orthogonal states of the field for which the polarization of the output is the same as the polarization of the corresponding

input. The discussion of polarization dispersion involves a different set of eigenstates, as we shall soon see.

Polarization dispersion

Polarization dispersion, or polarization mode dispersion, or simply PMD, describes the change with frequency of the polarization of the field at the output of a fiber while the input polarization is held constant. It is intimately connected with changes in the mean time delay of a pulse traversing the fiber as a function of the polarization of the input pulse.

For a fiber of length L, we have $\tilde{U}(L)=\mathbf{T}(L,0)\tilde{U}(0)$. Polarization dispersion derives from the frequency derivative of the Jones matrix $\mathbf{T}(L,0)$. We can write this in the general form

$$\frac{\partial}{\partial\omega}\mathbf{T}(L,0) = i(\tau_0\mathbf{I} + \frac{1}{2}\vec{\tau}\cdot\vec{\sigma})\,\mathbf{T}(L,0) \tag{5.1}$$

where again the Hermitian form of the operator in parentheses on the right is necessary to preserve the unitary nature of the Jones matrix as the frequency varies. Looking back at the form of the Jones matrix, Eq.(4.2a), it is important to realize that while τ_0 is the frequency derivative of θ_0, $\vec{\tau}$ is not the frequency derivative of $\vec{\theta}$, because $(\partial\vec{\theta}/\partial\omega)\cdot\vec{\sigma}$ does not commute with $\vec{\theta}\cdot\vec{\sigma}$. One can, however, evaluate $\vec{\tau}$ in terms of $\vec{\theta}$ by expanding the Jones matrix Eq.(4.2a) using Eq.(1.7d), differentiating with respect to frequency, and then rewriting the result in the form required by Eq.(5.1). The result is

$$\tau_0 = \frac{\partial\theta_0}{\partial\omega} \quad ; \quad \vec{\tau} = \frac{\partial\theta}{\partial\omega}\hat{\theta} + \sin\theta\frac{\partial\hat{\theta}}{\partial\omega} - (1-\cos\theta)\hat{\theta}\times\frac{\partial\hat{\theta}}{\partial\omega} \tag{5.2}$$

While this is an existence demonstration for $\vec{\tau}$, it does not seem to lead anywhere else. Note that τ_0 and $\vec{\tau}$ are frequency dependent quantities with dimensions of time.

To evaluate τ_0 and $\vec{\tau}$ in terms of the properties of the fiber, we recall that the Jones matrix can be expanded as an ordered product of the Jones matrices for each short section. For a section of differential length at position z, we have from Eq.(3.1)

$$\mathbf{T}(z+dz,z) = (1+i\beta_0(z)\,dz)\,\mathbf{I} + \frac{i}{2}\vec{\beta}(z)\cdot\vec{\boldsymbol{\sigma}}\,dz \tag{5.3}$$

Using the frequency derivative of Eq.(5.3), we get

$$\frac{\partial}{\partial\omega}\mathbf{T}(L,0) = i\int_0^L dz\,\mathbf{T}(L,z)\left[\frac{\partial\beta_0(z)}{\partial\omega} + \frac{1}{2}\frac{\partial\vec{\beta}(z)}{\partial\omega}\cdot\vec{\boldsymbol{\sigma}}\right]\mathbf{T}(z,0) \tag{5.4}$$

Rewriting $\mathbf{T}(z,0)$ as $\mathbf{T}(z,L)\,\mathbf{T}(L,0)$ and comparing the result with Eq.(5.1), we obtain

$$\tau_0 + \frac{1}{2}\vec{\tau}\cdot\vec{\boldsymbol{\sigma}} = \int_0^L dz\,\mathbf{T}(L,z)\left[\frac{\partial\beta_0(z)}{\partial\omega} + \frac{1}{2}\frac{\partial\vec{\beta}(z)}{\partial\omega}\cdot\vec{\boldsymbol{\sigma}}\right]\mathbf{T}(z,L) \tag{5.5}$$

Making use of Eq.(4.8), we finally arrive at the results

$$\tau_0 = \int_0^L dz\,\frac{\partial\beta_0(z)}{\partial\omega} \quad ; \quad \vec{\tau} = \int_0^L dz\,\overleftrightarrow{M}(L,z)\cdot\frac{\partial\vec{\beta}(z)}{\partial\omega} \tag{5.6}$$

Thus $\vec{\tau}$ is the frequency derivative of the birefringence vector $\vec{\beta}$, integrated over the length of the fiber, each element of the integral being projected to the end of the fiber via the appropriate Mueller matrix. Eq.(5.6) yields a simple rule for finding $\vec{\tau}$ for a fiber made of many sections. Each section contributes its own $\vec{\tau}$, projected to the end of the line by the Mueller matrix of the intervening fiber. The contributions from all sections are simply added together.

To observe the polarization dispersion, one varies the frequency of a monochromatic field entering a fiber of length L. Using Eq.(4.1), the frequency derivative of the Stokes vector at the output of the fiber is

$$\frac{\partial\vec{S}(L)}{\partial\omega} = \tilde{U}^\dagger(0)\,\mathbf{T}(0,L)\vec{\boldsymbol{\sigma}}\mathbf{T}(L,0)\frac{\partial\tilde{U}(0)}{\partial\omega} + \frac{\partial\tilde{U}^\dagger(0)}{\partial\omega}\mathbf{T}(0,L)\vec{\boldsymbol{\sigma}}\mathbf{T}(L,0)\,\tilde{U}(0)$$
$$+\ \tilde{U}^\dagger(0)\,\mathbf{T}(0,L)\left[\frac{i}{2}\vec{\boldsymbol{\sigma}}(\vec{\tau}\cdot\vec{\boldsymbol{\sigma}}) - \frac{i}{2}(\vec{\tau}\cdot\vec{\boldsymbol{\sigma}})\vec{\boldsymbol{\sigma}}\right]\mathbf{T}(L,0)\,\tilde{U}(0) \tag{5.7}$$

Using Eqs.(1.7a) and (4.6), this reduces to

$$\frac{\partial\vec{S}(L)}{\partial\omega} = \overleftrightarrow{M}(L,0)\cdot\frac{\partial\vec{S}(0)}{\partial\omega} - \vec{\tau}\times\vec{S}(L) \tag{5.8}$$

Experiments based on Eq.(5.8) are used to evaluate the magnitude of $\vec{\tau}$.

It is important to note that because of the vector addition in Eq.(5.6), $\vec{\tau}$ varies rapidly with frequency. Starting from Eq.(5.5), and neglecting the contribution of the $\partial^2\vec{\beta}/\partial\omega^2$, one can show that

$$\frac{\partial\vec{\tau}(L)}{\partial\omega} \approx \int_0^L dz\,\overleftrightarrow{M}(L,z)\cdot\left(\vec{\tau}(z)\times\frac{\partial\vec{\beta}(z)}{\partial\omega}\right) \tag{5.9}$$

where $\vec{\tau}(z)$ is the value of $\vec{\tau}$ for the piece of the fiber extending from 0 to z. In order to let $\vec{S}$ at the end of the fiber precess through a complete cycle according to Eq.(5.8), the frequency shift required is $2\pi/\tau$. Comparing Eq.(5.9) with Eq.(5.6), one might expect that for a long fiber this frequency shift is enough to produce a considerable change in $\vec{\tau}$, so that as the frequency changes, the Stokes vector wanders rather randomly over the Poincaré sphere.

An interesting result can be derived from Eq.(5.6). Taking the dot product of $\vec{\tau}$ with itself gives

$$\tau^2 = \int_0^L dz \int_0^L dz' \left[\overleftrightarrow{M}(L,z)\cdot\frac{\partial\vec{\beta}(z)}{\partial\omega}\right]\cdot\left[\overleftrightarrow{M}(L,z')\cdot\frac{\partial\vec{\beta}(z')}{\partial\omega}\right]$$

$$= \int_0^L dz \int_0^L dz' \,\frac{\partial\vec{\beta}(z)}{\partial\omega}\cdot\overleftrightarrow{M}(z,z')\cdot\frac{\partial\vec{\beta}(z')}{\partial\omega} \tag{5.10}$$

The integrand of Eq.(5.10) merits some discussion. If z and z' are close enough so that $\vec{\beta}$ can be considered constant, then $\overleftrightarrow{M}(z,z')$ represents a precession around $\vec{\beta}$. If $\partial\vec{\beta}/\partial\omega$ is directed along $\vec{\beta}$, as may usually be true, then the rotation does not matter, and the integrand reduces to $(\partial\beta/\partial\omega)^2$. For short lengths of fiber, then, one expects that τ^2 will grow with the square of the length. For much longer lengths, the integral over z', say, will roughly approach a constant, and so τ^2 will grow linearly with the length.

Time delays

To discuss time delays, one must think about pulses rather than monochromatic fields. At some position z, let the field of a pulse be represented in the time domain by $U(z,t)$, and in the frequency domain by $\tilde{U}(z,\omega)$. These two functions are Fourier transforms of each other. The tilde distinguishes between them, so that U is a function of time and $\tilde{U}$ is the corresponding function of angular frequency. Let the fields be normalized so that the energy in the pulse is

$$W = \int dt\, U^\dagger U = \int d\omega\, \tilde{U}^\dagger \tilde{U} \tag{6.1}$$

These integrals are complete, as are those in the following. Note that in this section, $\tilde{U}$ represents a monochromatic field, as before. Now, however, it is normalized so that $\tilde{U}^\dagger\tilde{U}$ represents the spectral energy density of the pulse.

The mean time at which a pulse passes some location z is defined by

$$\langle t(z)\rangle = \frac{1}{W}\int dt\, t\, U^\dagger(z)\, U(z) = \frac{1}{W}\int d\omega \tilde{U}^\dagger(z)\left[-i\frac{\partial \tilde{U}(z)}{\partial\omega}\right] \tag{6.2}$$

The second integral in Eq.(6.2) derives from the first by Fourier transform; it is real, as

can be shown by partial integration. Using $\tilde{U}(z) = \mathbf{T}(z,0)\,\tilde{U}(0)$, with Eq.(5.1) we get

$$\langle t(z)\rangle = \langle t(0)\rangle + \frac{1}{W}\int d\omega \tilde{U}^\dagger(z)\left[\tau_0(z)\mathbf{1} + \frac{1}{2}\vec{\tau}(z)\cdot\vec{\sigma}\right]\tilde{U}(z) \tag{6.3}$$

The first term of the integral defines the mean value of the frequency dependent quantity $\tau_0(z)$, whence Eq.(6.3) becomes

$$\langle t(z)\rangle = \langle t(0)\rangle + \langle\tau_0(z)\rangle + \frac{1}{2W}\int d\omega\, \vec{\tau}(z)\cdot\vec{S}(z) \tag{6.4}$$

where the length of $\vec{S}(z)$ is $\tilde{U}^\dagger(z)\tilde{U}(z)$, the spectral energy density of the pulse. In Eqs.(6.3) and (6.4), $\tau_0(z)$ is the propagation time delay of the fiber, neglecting the effects of birefringence, at frequency ω. The frequency average of $\tau_0(z)$ takes into account the effects of the ordinary time delay dispersion of the fiber. The averaging symbol $\langle\ \rangle$ generally means an average for the pulse. For a time dependent quantity, the average is most obviously expressed in the time domain, but as in Eq.(6.2) it can also be written in the frequency domain. For a frequency dependent quantity, the reverse is true. The final term represents the effects of birefringence. If the bandwidth is sufficiently narrow, and the field $\tilde{U}(z)$ is in one of the two eigenstates of $\vec{\tau}(z)\cdot\vec{\sigma}$, then $S(z)$ will be aligned either parallel or antiparallel to $\vec{\tau}(z)$, and so the last term will give an additional time delay of $\pm\tau(z)/2$. The difference, $\tau(z)$, is called the PMD time delay of the fiber, again at frequency ω. To get the average PMD time delay, the value of τ must be measured at a variety of frequencies, or over time as the propagation properties of the fiber slowly drift.

It is useful to project the vector dot product of Eq.(6.4) back to the input of the fiber, because while the input polarization is likely to be independent of frequency, the output polarization is not. Doing this we get

$$\langle t(z)\rangle = \langle t(0)\rangle + \langle\tau_0(z)\rangle + \frac{1}{2W}\int d\omega\,(\overleftrightarrow{M}(0,z)\cdot\vec{\tau}(z))\cdot\vec{S}(0) \tag{6.5}$$

For short pulses, approaching the PMD time delay $\tau(z)$, the direction of $\overleftrightarrow{M}(0,z)\cdot\vec{\tau}(z)$ will vary significantly over the bandwidth of the pulse. As a result, the whole pulse cannot be in an eigenstate of $\vec{\tau}(z)\cdot\vec{\sigma}$, and an increase in pulse duration will result.

Summary and Discussion

The purpose of this exercise is to show how the matrix-vector notation embodied in $\vec{\sigma}$ makes it possible to write the theory of propagation in lossless birefringent fibers in a compact form. The eigenstates of operators of the form $\vec{a}\cdot\vec{\sigma}$ play a central role. The intimate connection between Jones matrices and the Stokes space is immediately evident, for example Eq.(4.2a), where the vector $\vec{\theta}$ is a vector in Stokes space. The quantitative connection between polarization dispersion of the Stokes vector and the corresponding polarization dependent time delay is made rather easily. We have not included loss in this work. Polarization independent loss is no problem, since the theory does not include non-linearities, but obviously polarization dependent loss would change things.

Characteristic Units (t_c, z_c, P_c)

NLSE $\quad -i\dfrac{\partial u}{\partial z} = \dfrac{1}{2}\dfrac{\partial^2 u}{\partial t^2} + |u|^2 u$

$|u|^2 = \mathrm{sech}^2(t)$

$\tau = 1.763\, t_c$

Dispersion Length (from Linear Eqn.)

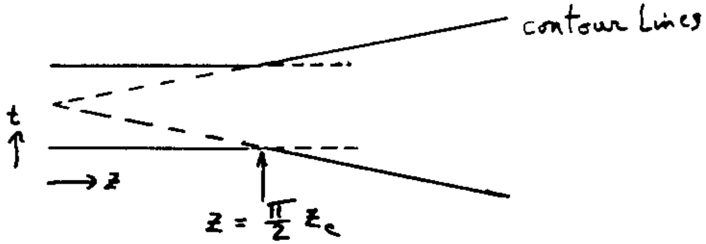
$$-i\frac{\partial u}{\partial z} = \frac{1}{2}\frac{\partial^2 u}{\partial t^2} \qquad k = -\frac{1}{2}\omega^2$$

$$\frac{\partial k}{\partial \omega} = v_g^{-1} = -\omega$$

$$u(t) \xrightarrow{F.T.} \tilde{u}(\omega) \xrightarrow[\text{far field}]{} u\left(\omega \to -\frac{t}{z}\right)$$

$$\mathrm{sech}\,(t) \longrightarrow \mathrm{sech}\left(\frac{\pi}{2}\omega\right) \longrightarrow \mathrm{sech}\left(\frac{\pi}{2}\frac{t}{z}\right)$$

asymtotic lines cross at $z = \pi/2$



at $\lambda = 1.555\ \mu m$ —

$$z_c\,(km) \cong \frac{1}{4}\,\frac{\tau\,(psec)^2}{D\,(psec/nm/km)}$$

τ	D	z_c
50	1	625
20	0.5	200

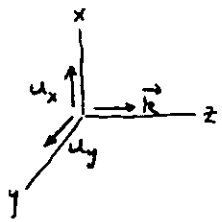
Canonical Soliton:

$$u = \operatorname{sech} t \, \exp\,(i z/2)$$

For distances $\ll z_c$, dispersion and nonlinearity are not very important!

2-DIM.

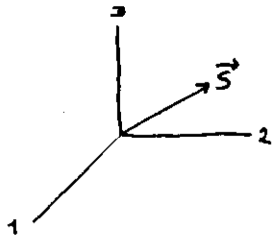
Field



$$U = \begin{pmatrix} u_x(z,t) \\ u_y(z,t) \end{pmatrix}$$

$$\tilde{U} = \begin{pmatrix} \tilde{u}_x(z,\omega) \\ \tilde{u}_y(z,\omega) \end{pmatrix}$$

Stokes Vector



$$S_1 = |u_x|^2 - |u_y|^2$$

$$S_2 = u_x^* u_y + u_x u_y^*$$

$$S_3 = -i u_x^* u_y + i u_x u_y^*$$

$$S_0 = |u_x|^2 + |u_y|^2 = U^\dagger U = \|\vec{S}\|$$

Pauli Matrices (re-ordered)

$$\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

with unit matrix; complete set of Hermitian 2x2 matrices

$$S_i = U^\dagger \sigma_i U$$

In vector notation

$$\vec{S} = U^{\dagger} \vec{\sigma}\, U$$

Think of $\vec{\sigma}$ as a vector in Stokes space

$$\vec{\sigma} = \sigma_1 \hat{e}_1 + \sigma_2 \hat{e}_2 + \sigma_3 \hat{e}_3$$

Properties of the σ matrices

$$\sigma_i^2 = \mathbb{I} \quad \text{(unit matrix)}$$

$$\sigma_1 \sigma_2 = -\sigma_2 \sigma_1 = i\,\sigma_3 \quad , \quad \text{etc.}$$

If $\vec{a}$ is any vector in Stokes space

$$\vec{a} = a_1 \hat{e}_1 + a_2 \hat{e}_2 + a_3 \hat{e}_3$$

then
$$\vec{a}\cdot\vec{\sigma} = a_1\sigma_1 + a_2\sigma_2 + a_3\sigma_3 = \begin{pmatrix} a_1 & a_2 - i a_3 \\ a_2 + i a_3 & -a_1 \end{pmatrix}$$

$$(\vec{a}\cdot\vec{\sigma})^2 = a^2 \qquad \vec{a} = a\,\underbrace{\hat{a}}_{\text{unit vector}}$$

$$\vec{\sigma}\,(\vec{a}\cdot\vec{\sigma}) = \vec{a} + i\,\vec{a}\times\vec{\sigma}$$

etc.

Birefringence

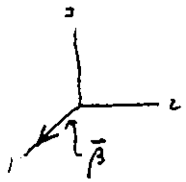
$$\frac{\partial \hat{U}}{\partial z} = i \left[\beta_0 \hat{U} + \tfrac{1}{2} \underbrace{\vec{\beta}\cdot\vec{\sigma}}_{\text{eigenvalues } \pm\beta} \hat{U} \right]$$

Hermitian

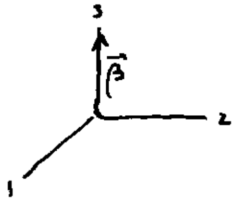
$$\frac{\partial \vec{S}}{\partial z} = \tilde{U}^{\dagger} \vec{\sigma} \;\tfrac{i}{2}\, \vec{\beta}\cdot\vec{\sigma}\, \tilde{U} \quad + \text{h.c.}$$

$$= -\tilde{U}^{\dagger}\, \vec{\beta}\times\vec{\sigma}\, \tilde{U}$$

$$= -\vec{\beta}\times\vec{S} \qquad [= 0 \text{ for eigenstates}]$$



= linear birefringence



= circular birefringence

β = magnitude of birefringence

Non-Linear Term

$$\frac{\partial U}{\partial z} = i\left[(U^\dagger U)U - \frac{1}{3}(U^\dagger \sigma_3 U)\,\sigma_3 U \right]$$

$$= i\left[S_0 U - \frac{1}{3} S_3 \sigma_3 U \right]$$

$$\sigma_3 = \hat{e}_3 \cdot \vec{\sigma}$$

$$\frac{\partial \vec{S}}{\partial z} = \frac{2}{3} S_3 \,(\hat{e}_3 \times \vec{S})$$

$$(U^\dagger \sigma_3 U)\,\sigma_3 U = \begin{pmatrix} |u_y|^2 u_x - u_y^2 u_x^* \\ |u_x|^2 u_y - u_x^2 u_y^* \end{pmatrix}$$

Identity $\quad \sum S_j \sigma_j U = S_0 U$

for any U

Propagation Eqn.

$$-i\frac{\partial U}{\partial z} = \frac{1}{2}\overset{\vec\beta(z)}{\vec{\beta}}\cdot\vec{\sigma}\,U + \frac{i}{2}\overset{\frac{\partial\vec\beta}{\partial\omega}}{\vec{\delta}}\cdot\vec{\sigma}\,\frac{\partial U}{\partial t}$$

$$+ \frac{1}{2}\frac{\partial^2 U}{\partial t^2} + (U^\dagger U)U \qquad \text{(manakov Eq.)}$$

$$+ \frac{1}{8}\left[(U^\dagger U)U - 3(U^\dagger \sigma_3 U)\,\sigma_3 U \right] \leftarrow \text{small?}$$

(NL term renormalized by 9/8)

Eliminate the $\vec{\beta}\cdot\vec{\sigma}$ term

Transform $\quad U(z,t) = T(z,0)\,U'(z,t)$

where the unitary (Jones) matrix $T(z,z')$ satisfies

$$-i\frac{\partial}{\partial z}T(z,z') = \frac{1}{2}\vec{\beta}(z)\cdot\vec{\sigma}\;T(z,z')$$

with $T(z,z) = 1$

Note: for constant $\vec{\beta}$, $\quad T(z,0) = \exp\left(\frac{iz}{2}\vec{\beta}\cdot\vec{\sigma}\right)$

$$T(z,0) = T(z,z')\,T(z',0)$$

Transformed NLSE

Set $U(z) = T(z,0)\,U'(z)$

corr. $\vec{S}(z) = \overleftrightarrow{M}(z,0)\cdot\vec{S}'(z)$

↖ Mueller matrix (3×3)

drop the primes on U

$$-i\frac{\partial U}{\partial z} = \frac{1}{2}\frac{\partial^2 U}{\partial t^2} + (U^\dagger U)U \quad \leftarrow \text{soliton}$$

$$+\frac{i}{2}(\vec{\delta}'\cdot\vec{\sigma})\frac{\partial U}{\partial t} \quad \leftarrow \text{polarization dispersion}$$

$$+\frac{1}{8}\left[(U^\dagger U)U - 3(U^\dagger\sigma_3' U)\sigma_3' U\right]$$

negligible?

$$\vec{\delta}' = \overleftrightarrow{M}(0,z)\cdot\vec{\delta}(z)$$

$$\sigma_3' = \left(\overleftrightarrow{M}(0,z)\cdot\hat{e}_3\right)\cdot\vec{\sigma}$$

We have got the Manakov eqn. + perturbations.

Solitons and Polarization Dispersion

$$\frac{\partial U}{\partial z} = -\frac{1}{2}\vec{\delta}'\cdot\vec{\sigma}\,\frac{\partial U}{\partial t}$$

$$U = U_s = U_0\,\mathrm{sech}(t-t_m)$$

U_0 = normalized polarization state $U_0^\dagger U_0 = 1$

Set $U_{0\perp}$ be a polarization state orthogonal to U_0

[e.g. $U_0 = \begin{pmatrix}1\\0\end{pmatrix}$, $U_{0\perp} = \begin{pmatrix}0\\1\end{pmatrix}$]

then $U_{0\perp}^\dagger U_0 = 0$, and $U_0 U_0^\dagger + U_{0\perp}U_{0\perp}^\dagger = I$

completeness

We have then

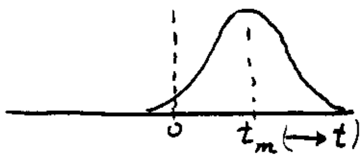
$$\vec{\sigma}\,\frac{\partial U_s}{\partial t} = U_0\left(U_0^\dagger\vec{\sigma}\,\frac{\partial U_s}{\partial t}\right) + U_{0\perp}\left(U_{0\perp}^\dagger\vec{\sigma}\,\frac{\partial U_s}{\partial t}\right)$$

time delay   loss

$$\frac{\partial U_s}{\partial z} = -\frac{1}{2}\vec{\delta}'\cdot\hat{S}_0\,\frac{\partial U_s}{\partial t} \qquad \hat{S}_0 = U_0^\dagger\vec{\sigma}\,U_0$$

$$\frac{dt_m}{dz} = \frac{1}{2}\vec{\delta}'\cdot\hat{S}_0$$

PMD time delay



$$\frac{dt_m}{dz} = \frac{1}{2}\,\vec{\delta}'\cdot\hat{S}_0$$

normalized (unit) Stokes vector of U

$$\tau_m = \frac{1}{2}\,\vec{\tau}(L)\cdot\hat{S}_0$$

$$\vec{\tau}(L) = \int_0^L dz\,\overbrace{\overleftrightarrow{M}(0,z)\cdot\vec{\delta}(z)}^{\vec{\delta}'}$$

$$\langle\|\vec{\tau}(L)\|^2\rangle = \int_0^L dz\int_0^L dz'\left\langle\vec{\delta}(z')\cdot\overleftrightarrow{M}(z',z)\cdot\vec{\delta}(z)\right\rangle$$

$$\cong\langle\delta^2\rangle\,\ell_{coh}\,L$$

for good fibers $\tau_{rms} \cong 0.1$ psec/km$^{1/2}$

$\delta \sim .7$ psec/km ; $\ell_{coh} \sim 20$ m.

Perturbations of Solitons of the Manakov Equation

$$-i\frac{\partial U}{\partial z} = \frac{1}{2}\frac{\partial^2 U}{\partial t^2} + (U^\dagger U)U + \text{perturbations}$$

A. In the same polarization as the soliton, we are back to one dimension

$$U = \begin{pmatrix} u_s + \epsilon u_p \\ 0 \end{pmatrix}$$

modification of soliton parameters + radiation

B. In the orthogonal polarization

$$U = \begin{pmatrix} u_s \\ \epsilon u_p \end{pmatrix}$$

modification of the soliton polarization + radiation.

C. By collisions between solitons.

modification of the soliton's polarization

Perturbations in 1 dimension

$$\frac{\partial u}{\partial z} = i\left[\frac{1}{2}\frac{\partial^2 u}{\partial t^2} + u^* u^2\right] + R(z,t)\, u$$

canonical soliton

$$u_s = \operatorname{sech} t \exp(iz/2)$$

[wavenumber $k = 1/2$]

small linear waves

$$u = \epsilon \exp(ikz - i\omega t)$$

[wavenumber $k = -\frac{1}{2}\omega^2$]

a perturbation with wavenumber $-k_R$

couples power from the soliton at a frequency

given by $\quad k_R = \frac{1}{2}(1+\omega^2)$

or $\quad \omega = \sqrt{2k_R - 1}$

$$-i\frac{\partial u}{\partial z} = \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u$$

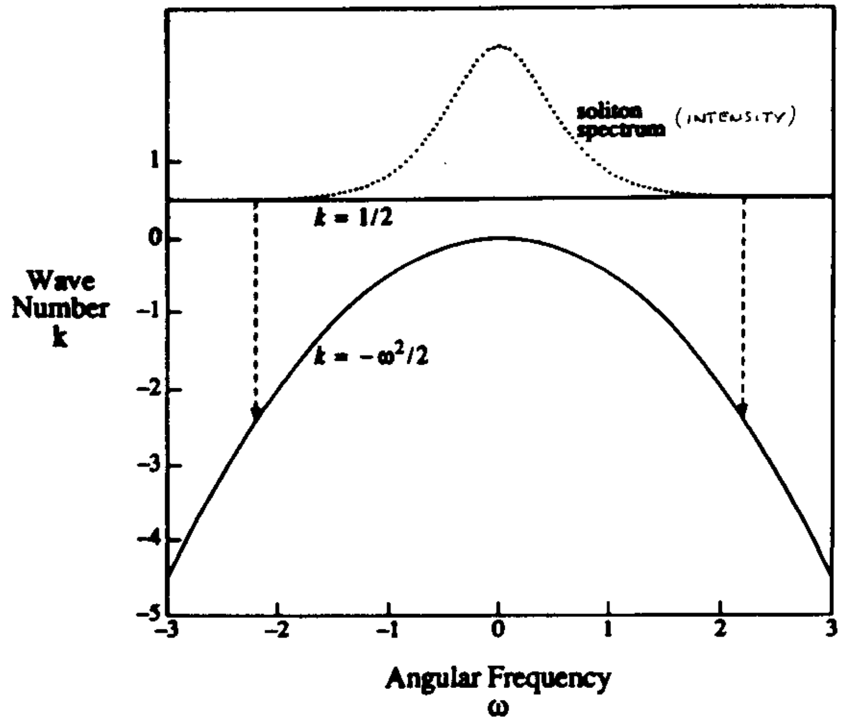
Low power dispersive waves

$$u \propto e^{ikz - i\omega t} \qquad k = -\frac{1}{2}\omega^2$$

Soliton: $u = \operatorname{sech} t\, e^{iz/2} \qquad k = \frac{1}{2}$

To couple power out, a perturbation $\propto \cos(Kz)$

sees only the frequency $\omega = \sqrt{2K-1}\ (\pm)$



Inverse Scattering Transform for radiative perturbations

$$\frac{\partial u}{\partial z} = i\left[\frac{1}{2}\frac{\partial^2 u}{\partial t^2} + u^* u^2\right]$$

$u = u_s + |\epsilon|\, u_p \quad ; \quad u_s = \operatorname{sech} t \exp(iz/2)$

$$\frac{\partial u_p}{\partial z} = i\left[\frac{1}{2}\frac{\partial^2 u_p}{\partial t^2} + 2|u_s|^2 u_p + u_s^2 u_p^*\right] \qquad (1)$$

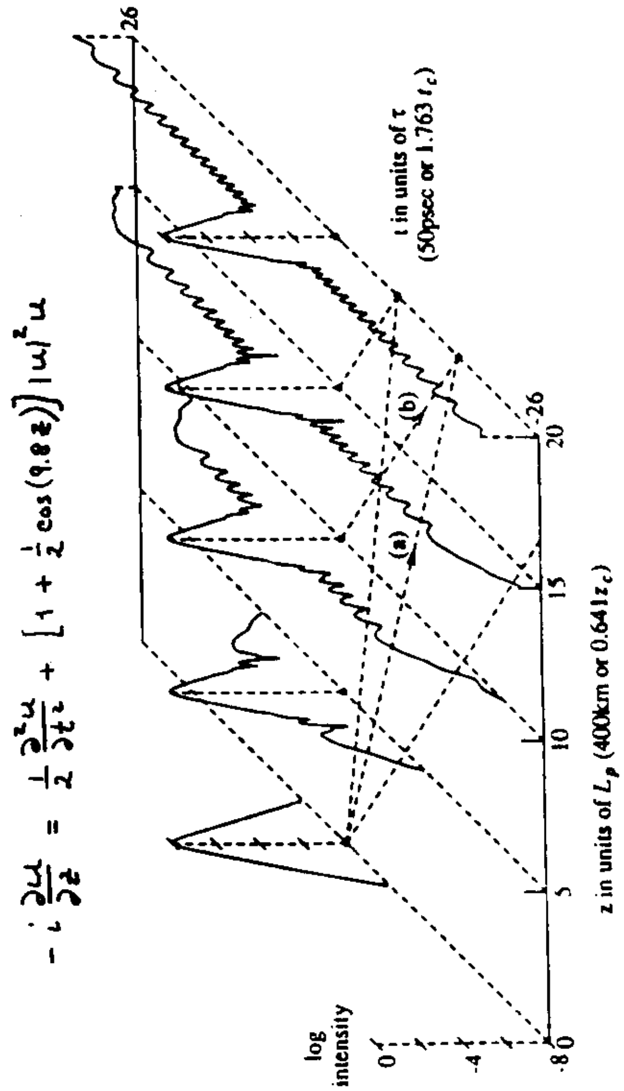
If $f(z,t)$ satisfies the linear equation

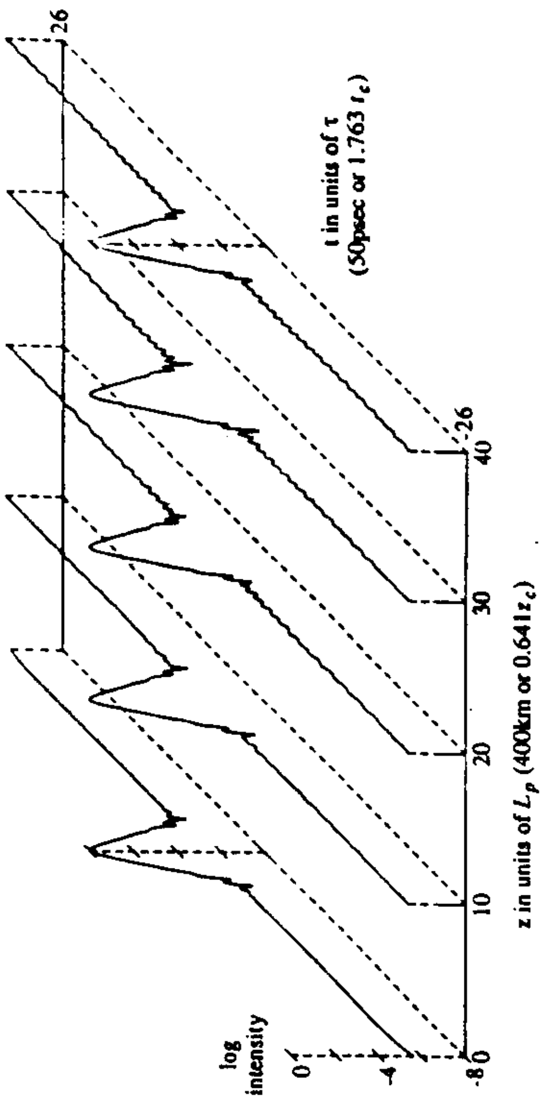
$$\frac{\partial f}{\partial z} = i\,\frac{1}{2}\frac{\partial^2 f}{\partial t^2}$$

then

$$u_p = -\frac{\partial^2 f}{\partial t^2} + 2\tanh t\,\frac{\partial f}{\partial t} - (\tanh^2 t) f + u_s^2 f^*$$

satisfies Eq. (1)





Manakov Eqn.

Let u, v be orthogonal polarizations

$$-i\frac{\partial u}{\partial z} = \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + (|v|^2 + |u|^2)\,u$$

$$-i\frac{\partial v}{\partial z} = \frac{1}{2}\frac{\partial^2 v}{\partial t^2} + (|v|^2 + |u|^2)\,v$$

solitons have $|u|^2 + |v|^2 = \operatorname{sech}^2 t$

Let u be the soliton and $|v|^2 \lll |u|^2$

$$-i\frac{\partial u}{\partial z} = \frac{1}{2}\frac{\partial^2 u}{\partial t^2} + |u|^2 u$$

$$u = \operatorname{sech}(t)\, e^{iz/2}$$

$$-i\frac{\partial v}{\partial z} = \frac{1}{2}\frac{\partial^2 v}{\partial t^2} + (\operatorname{sech}^2 t)\,v$$

solutions?

$$-i\frac{\partial v}{\partial z} = \frac{1}{2}\frac{\partial^2 v}{\partial t^2} + (\operatorname{sech}^2 t)\, v$$

KdeV scattering problem

$$v = \epsilon \operatorname{sech} t \exp(iz/2)$$

(bound sol'n ; changes polarization of field)

$$v = \epsilon(\tanh t + i\omega)\exp(ikz - i\omega t)$$

$$k = -\frac{1}{2}\omega^2$$

$$v = \epsilon\left(\tanh t - \frac{\partial}{\partial t}\right)\exp(ikz - i\omega t)$$

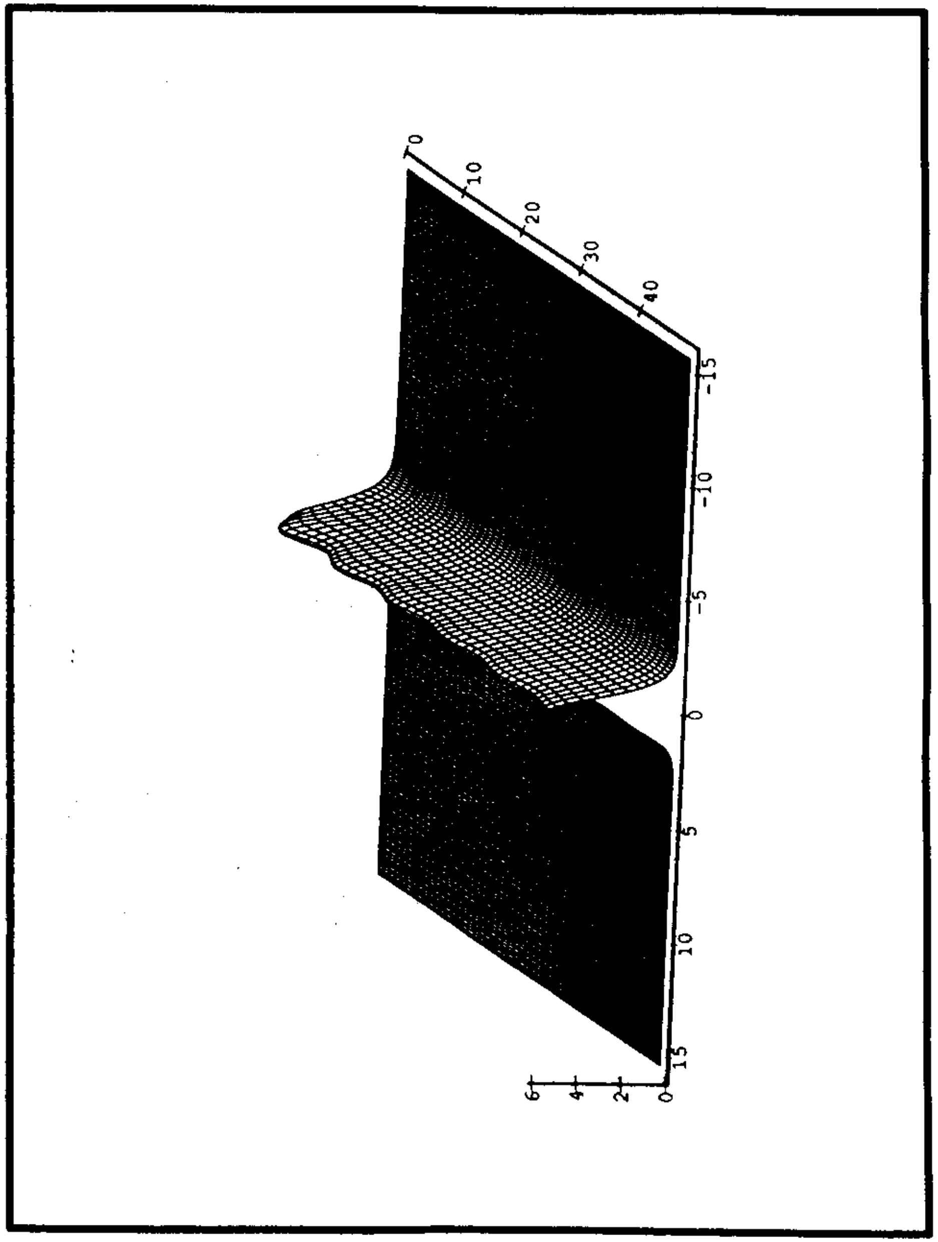
$$v = \epsilon\left(\tanh t - \frac{\partial}{\partial t}\right) f(z,t)$$

where $\quad -i\dfrac{\partial f}{\partial z} = \dfrac{1}{2}\dfrac{\partial^2 f}{\partial t^2}$

at some z,

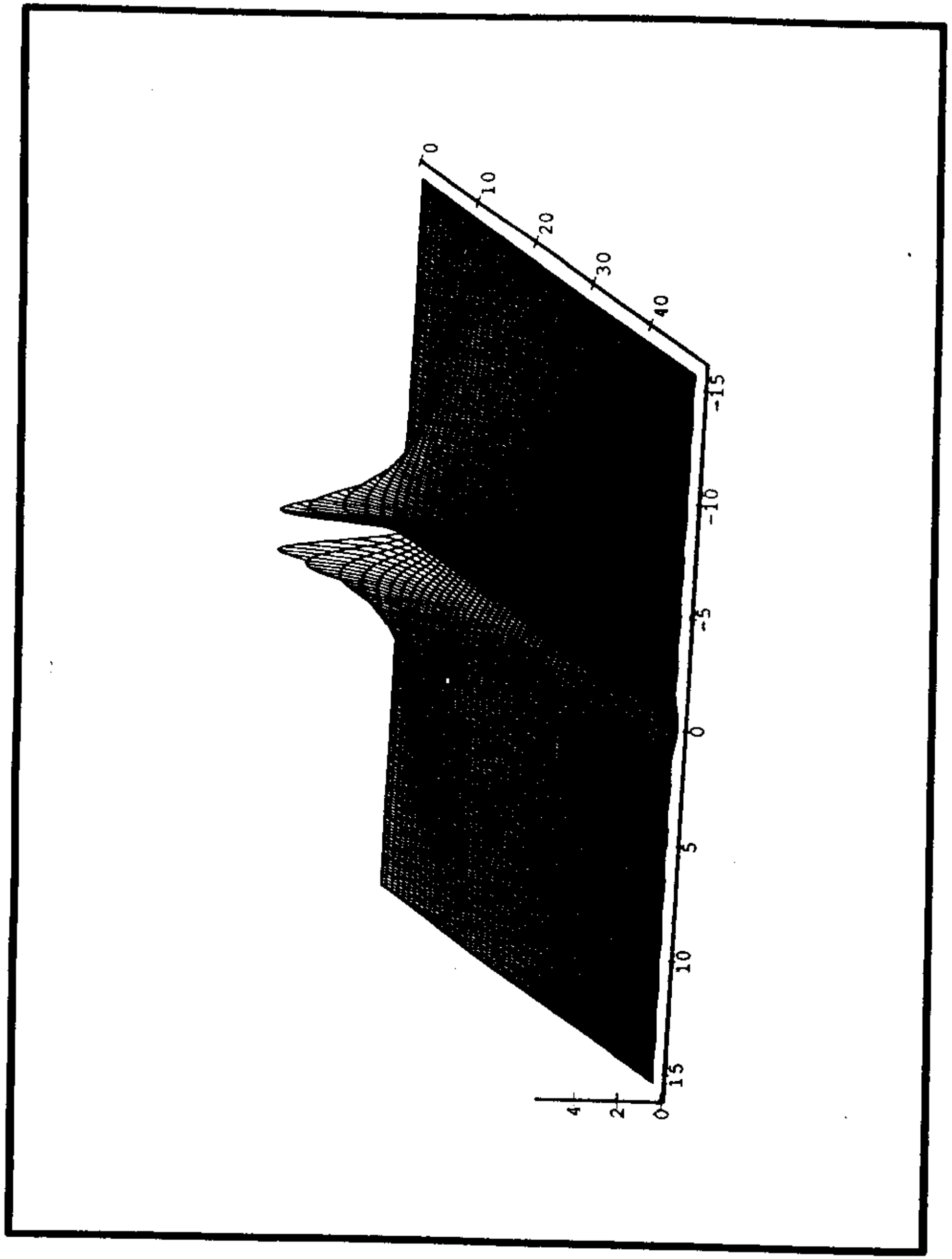
$$f = \operatorname{sech} t \quad ; \quad v = 2\epsilon \tanh t \operatorname{sech} t$$

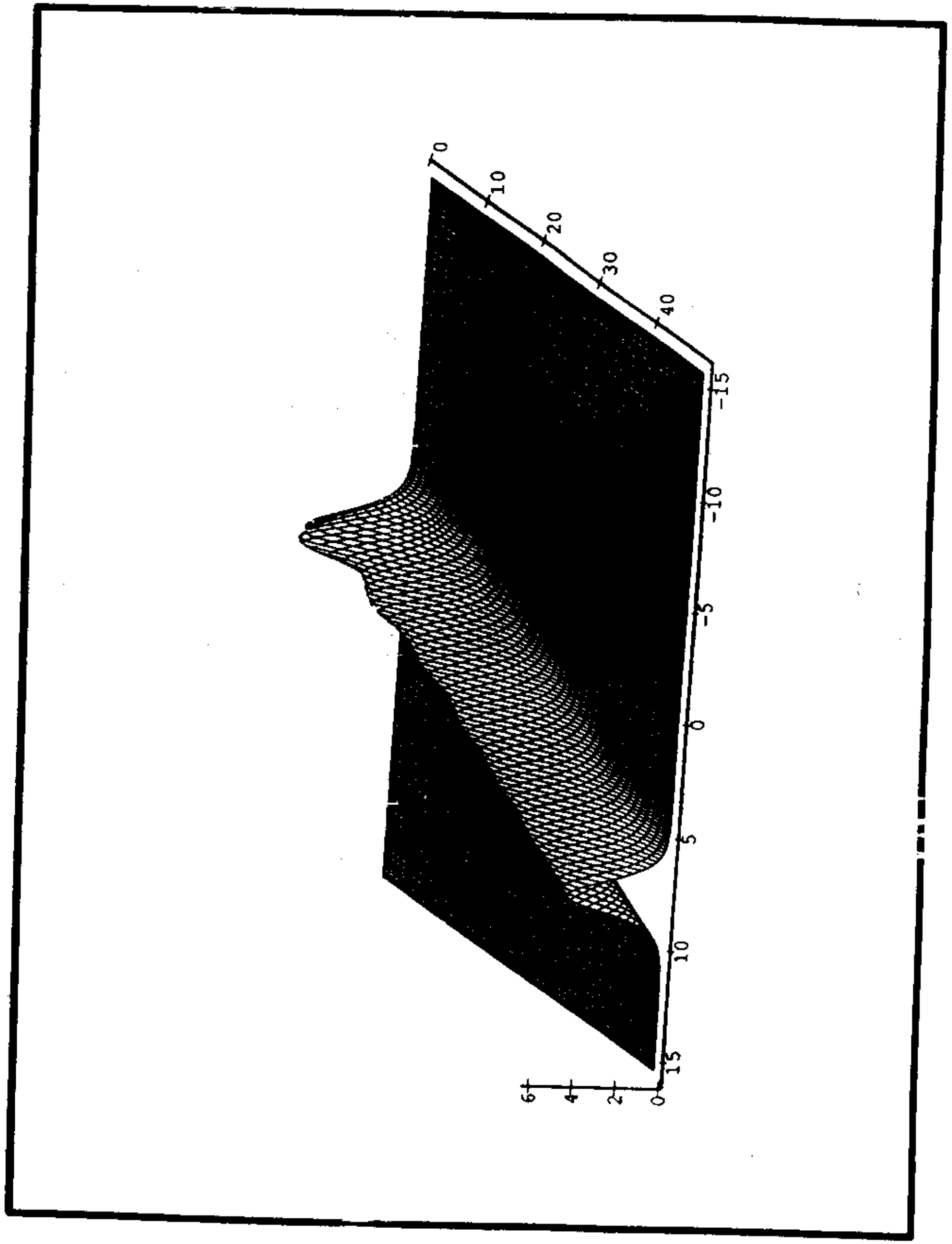
(example)

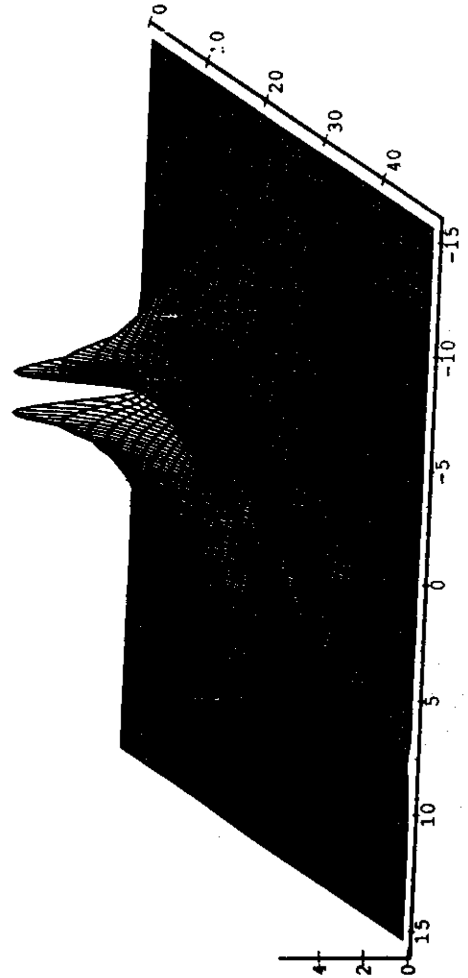


20

36







Soliton Collisions

$$-i\frac{\partial U}{\partial z} = (U^{\dagger}U)U$$

If $U = U_{s1} + U_{s2}$ (solitons of different frequencies)

$$-i\frac{\partial U_{s1}}{\partial z} = \underbrace{(U_{s1}^{\dagger}U_{s1})U_{s1}}_{\text{self phase mod.}} + \underbrace{(U_{s2}^{\dagger}U_{s2})U_{s1}}_{\text{cross phase mod.}} + (U_{s2}^{\dagger}U_{s1})U_{s2} \qquad (!)$$

$$U_{s1} = U_1 \operatorname{sech} t \quad ; \quad U_{s2} = U_2 \operatorname{sech}(t+\Omega z)$$

$$\text{collision length} = \frac{2\times 1.763}{\Omega} < 1$$

Neglecting effects of small phase shifts, from (!) we get

$$\Delta U_1 \cong i\,(U_2^{\dagger}U_1)U_2 \int dz\, \operatorname{sech}^2(t+\Omega z)$$

$$= i\,\frac{2}{\Omega}(U_2^{\dagger}U_1)U_2$$

One can show $\Delta\hat{S}_1 \cong \frac{2}{\Omega}(\hat{S}_1 \times \hat{S}_2)$