Budden & Smith's 'additional memory' and the geometric phase

By M. V. Berry

H. H. Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, U.K.

Budden & Smith considered vector waves in a z-stratified medium, driven by an arbitrary matrix $\mathsf{A}(X)$ depending on parameters $X(z) = \{X_1(z), X_2(z), \ldots\}$ which characterize the medium. They showed that in the short-wave limit there is not only the familiar optical path phase factor but an 'additional memory' factor M, whose exponent is the line integral, along the ray, of a certain 1-form constructed from the eigenvectors of A. Their discovery anticipated the geometric phase of quantum mechanics (where z is time, and A is the hermitian hamiltonian operator). Explicit connection is made between the two formalisms. If A is symmetric, or can be made symmetric by multiplication by a constant matrix, the 1-form is integrable and M does not represent memory because it can be expressed locally, in terms of the properties of the medium at the endpoints of the ray; in quantum mechanics, symmetrizability is equivalent to the hamiltonian possessing antiunitary symmetry (e.g. time reversal). A non-integrable real M arises when light traverses a transparent medium with variable refractive index and optical activity. However, the non-integrability here is cancelled by an extra contribution from the ordinary optical path length, arising from an additional term in the constitutive relation between electric field and displacement, which is necessary in an inhomogeneous medium to ensure its transparency.

1. Introduction

Smith (1975) and Budden & Smith (1976, hereafter called BS) described a novel aspect of vector wave propagation in stratified media. This emerges in the asymptotic limit of short wavelength, where the appropriate description is in terms of rays. Not only is there the familiar 'phase memory', namely the non-local dependence of the wave on the properties of the medium, embodied in the complex exponential whose phase is the optical path length (integral of the local wavenumber along the ray). There is, in general, a further exponential factor which is also non-local in the sense that it involves an inexact differential which cannot be integrated so as to depend only on the medium at the endpoints of the ray. This 'additional memory' can be real or complex, and so can be an amplitude as well as a phase.

My reasons for returning to this subject now are fourfold. First, to draw attention to Budden & Smith's little-known discovery, which with hindsight can be regarded as one of several substantial anticipations of the geometric phase of quantum mechanics (Berry 1984; Shapere & Wilczek 1989; Markovski & Vinitsky 1989; Berry 1991); 'additional memory' corresponds to the geometric phase, and 'phase memory' to the dynamical phase. Secondly, to make explicit the connection between the formalism of BS and that of the geometric phase, especially its non-hermitian generalization by Garrison & Wright (1988). Thirdly, to simplify and generalize an interesting condition given by BS, which guarantees the vanishing of the additional memory. And fourthly, to expose an instructive error in one of the examples given by BS, which illustrates how the existence of additional memory can reveal unphysical features in wave equations.

2. Memory formalism

BS consider a monochromatic wave with wavenumber k, whose non-trivial space dependence is confined to a single coordinate z and which is represented by an N-component vector $\boldsymbol{e}(z)$. The stratified propagation medium is represented by a set of parameters $X(z) = \{X_1(z), X_2(z), \ldots\}$. These are incorporated into an $N \times N$ matrix $\mathsf{A}(X)$ which governs the wave according to

$$(\mathrm{i}/k)\,\boldsymbol{e}' = \mathsf{A}\boldsymbol{e}, \tag{1}$$

where the prime denotes $\mathrm{d}/\mathrm{d}z$ and $\boldsymbol{e}$ is represented by a column matrix. The matrix A need not be real, symmetric, or hermitian. As BS point out, a great variety of wave equations can be expressed in this form; their examples include electromagnetic waves in plasmas and dielectrics, electroacoustic waves, atmospheric gravity waves, seismic waves, and magnetohydrodynamic waves. Quantum mechanics is included too, if z is identified as time, $\boldsymbol{e}$ as the evolving state (vector in Hilbert space), k as the reciprocal of Planck's constant, and A as the (hermitian) driving hamiltonian.

The régime in which the wavelength is short (relative to the variation of the medium parameters X) corresponds to large k. In quantum mechanics this is the adiabatic régime. According to the phase-integral or WKB method, the wave $\boldsymbol{e}$ clings to the local eigenvectors (modes) $\boldsymbol{u}_n(X)$, which satisfy

$$\mathsf{A}\boldsymbol{u}_n = q_n \boldsymbol{u}_n, \tag{2}$$

where q_n denotes the eigenvalue (assumed non-degenerate) corresponding to $\boldsymbol{u}_n$. It is convenient in calculations not to impose any normalization on $\boldsymbol{u}_n$. BS show that the wave with initial condition

$$\boldsymbol{e}(0) = \boldsymbol{u}_n[X(0)] \tag{3}$$

propagates to

$$\boldsymbol{e}(z) = \exp\left\{\int_0^z \mathrm{d}z\,\Gamma_n(z)\right\}\exp\left\{-\mathrm{i}k\int_0^z \mathrm{d}z\,q_n(z)\right\}\boldsymbol{u}_n[X(z)]. \tag{4}$$

In this formula, the third factor gives the form of the mode, and depends only on the local properties of the medium at z. The second factor is the ordinary phase memory. The first factor, to be denoted by M, is the additional memory which is our concern here. BS obtain

$$\Gamma_n = -\left(\mathsf{S}^{-1}\frac{\mathrm{d}}{\mathrm{d}z}\mathsf{S}\right)_{nn}, \tag{5}$$

where $\mathsf{S}\,(=\mathsf{S}(X))$ is the $N \times N$ matrix whose columns are the N eigenvectors $\boldsymbol{u}_n(X)$. Indexing the components of $\boldsymbol{u}_n$ by μ, the elements of S are

$$\mathsf{S}_{\mu n} = u_{n\mu}. \tag{6}$$

To make connections with more recent work, we first note that in (4) and (5) the dzs cancel in the integral for additional memory, so that this depends only on the path $X(z)$ through the space of medium parameters X. Alternatively stated, additional memory is geometric memory. In particular, it is independent of the wavenumber k. There may be a local part (which as we shall illustrate later 'cures' any inappropriate normalization for $\boldsymbol{u}_n$) but this can be eliminated by choosing the path in X space to be a circuit, so that the medium at the endpoints is the same. For circuits, any additional memory is entirely non-local.

Next, we express S^{-1} in terms of the left eigenvectors of A, that is the row vectors $\boldsymbol{v}_n^{\mathrm{T}}$ formed by transposition of the eigenvectors $\boldsymbol{v}_n$ of A^{T}:

$$\boldsymbol{v}_n^{\mathrm{T}}\mathsf{A} = q_n\,\boldsymbol{v}_n^{\mathrm{T}}, \quad \text{where} \quad \mathsf{A}^{\mathrm{T}}\boldsymbol{v}_n = q_n\,\boldsymbol{v}_n. \tag{7}$$

From biorthogonality it follows that

$$(\mathsf{S}^{-1})_{n\mu} = v_{n\mu}/\boldsymbol{v}_n\cdot\boldsymbol{u}_n, \tag{8}$$

where the dot denotes the ordinary scalar product of vectors, even when these are complex (this is not the same as the complex product $\boldsymbol{v}_n^*\cdot\boldsymbol{u}_n$). Thus the additional memory factor can be written as

$$M = \exp\left\{-\int_{\text{path}}\frac{\boldsymbol{v}_n\cdot\mathrm{d}\boldsymbol{u}_n}{\boldsymbol{v}_n\cdot\boldsymbol{u}_n}\right\}, \tag{9}$$

where in the 1-form the differential d acts in X space.

This explicit expression for the additional memory, in terms of the local eigenvector and its rate of change, is the same as that given by Garrison & Wright (1988) in their generalization of the geometric phase to dissipative systems. Note that Garrison & Wright's formulae look different because they employ normalized eigenvectors, and also a complex scalar product and the hermitian conjugate (adjoint) rather than the transpose; I do not do this here because it obscures a simple integrability condition, to be discussed in the next section.

To obtain the familiar geometric phase of quantum mechanics, we observe that when A is hermitian, $\mathsf{A}^{\mathrm{T}} = \mathsf{A}^*$ and the left eigenvector defined by (7) is the row vector whose elements are the complex conjugates of those of the right eigenvectors. Then with the usual normalization $\boldsymbol{u}_n^*\cdot\boldsymbol{u}_n = 1$ the differential in (9) is imaginary, giving the phase factor found by Berry (1984):

$$M = \exp\left\{-\mathrm{i}\,\mathrm{Im}\int_{\text{path}}\boldsymbol{u}_n^*\cdot\mathrm{d}\boldsymbol{u}_n\right\}. \tag{10}$$

This geometric phase factor multiplies the ordinary 'dynamical' phase factor; in BS this is the ordinary phase memory, given by the second factor in (4).

3. Integrability conditions

If the right and left eigenvectors $\boldsymbol{u}_n$ and $\boldsymbol{v}_n^{\mathrm{T}}$ are known in the relevant region of parameter space X, the most immediate test for the integrability of the memory M is the vanishing of the corresponding 2-form, that is

$$\mathrm{d}\wedge\frac{\boldsymbol{v}_n\cdot\mathrm{d}\boldsymbol{u}_n}{\boldsymbol{v}_n\cdot\boldsymbol{u}_n} = \frac{\mathrm{d}\boldsymbol{v}_n\wedge\cdot\,\mathrm{d}\boldsymbol{u}_n}{\boldsymbol{v}_n\cdot\boldsymbol{u}_n} + \frac{\boldsymbol{v}_n\cdot\mathrm{d}\boldsymbol{u}_n\wedge\mathrm{d}\boldsymbol{v}_n\cdot\boldsymbol{u}_n}{(\boldsymbol{v}_n\cdot\boldsymbol{u}_n)^2} = 0. \tag{11}$$

It is, however, helpful to have tests that can be carried out directly on the medium matrix $\mathsf{A}(X)$.

A sufficient and rather general condition for integrability is that A is symmetric, or can be made symmetric by a similarity transformation involving a constant matrix R, for all $X(z)$ on the ray path. Thus if

$$A_R \equiv R^{-1}AR \tag{12}$$

is symmetric, then M is integrable. To show this, rewrite the fundamental equation (1) as

$$(\mathrm{i}/k)\,\boldsymbol{e}_R' = A_R\,\boldsymbol{e}_R, \quad \text{where} \quad \boldsymbol{e}_R \equiv R^{-1}\boldsymbol{e}. \tag{13}$$

Because A_R is symmetric, each of its left eigenvectors is simply the transpose of the corresponding right eigenvector, and the 1-form in (9) becomes

$$\frac{\boldsymbol{u}_{Rn}\cdot \mathrm{d}\boldsymbol{u}_{Rn}}{\boldsymbol{u}_{Rn}\cdot \boldsymbol{u}_{Rn}} = \frac{1}{2}\frac{\mathrm{d}(\boldsymbol{u}_{Rn}\cdot \boldsymbol{u}_{Rn})}{\boldsymbol{u}_{Rn}\cdot \boldsymbol{u}_{Rn}} = \mathrm{d}\ln\{\sqrt{(\boldsymbol{u}_{Rn}\cdot \boldsymbol{u}_{Rn})}\}. \tag{14}$$

Thus M becomes the purely local factor

$$M = M(z) = \sqrt{\frac{(\boldsymbol{u}_{Rn}\cdot \boldsymbol{u}_{Rn})_0}{(\boldsymbol{u}_{Rn}\cdot \boldsymbol{u}_{Rn})_z}} \tag{15}$$

involving the normalization of the eigenvectors. Note that neither the symmetrized matrix A_R nor its symmetrizer R need be real.

It is instructive to present this condition in a slightly different way. If A_R is symmetric, then, from (12),

$$R^{-1}AR = R^{\mathrm{T}}A^{\mathrm{T}}(R^{\mathrm{T}})^{-1}. \tag{16}$$

Thus

$$CA \text{ is symmetric, where } C = (RR^{\mathrm{T}})^{-1}. \tag{17}$$

But C is symmetric, and it follows that M is integrable if A can be symmetrized by multiplying it by a constant symmetric matrix.

This form of the condition gives an immediate proof of the result (BS) that M is integrable if A has symmetry about its trailing diagonal. Simply choose C to be the 'trailing unit matrix'

$$C = \begin{pmatrix} 0 & \cdot & \cdot & & 1 \\ \cdot & & & 1 & \\ \cdot & & 1 & & \cdot \\ & 1 & & & \cdot \\ 1 & & \cdot & \cdot & 0 \end{pmatrix} \tag{18}$$

and then CA is symmetric in the ordinary (leading diagonal) sense.

In quantum mechanics, the symmetrizability conditions for integrability are equivalent to requiring A to possess antiunitary symmetry (Porter 1965). To show this, the first step is to note that the driving matrix A in (1) must be hermitian, and if it is to remain hermitian after similarity transformations these must be induced by unitary matrices: $R = (R^{\dagger})^{-1}$. Thus the result A_R of the transformation (equation (12)) must be hermitian as well as symmetric, and therefore real symmetric. This implies

$$A_R = R^{\dagger}AR = A_R^{*} = R^{\mathrm{T}}A^{*}(R^{\mathrm{T}})^{-1} \tag{19}$$

that is

$$A = UA^{*}U^{\dagger}, \quad \text{where} \quad U = RR^{\mathrm{T}}. \tag{20}$$

Thus A is invariant under the transformation induced by $\boldsymbol{U}$ (a unitary operator) followed by complex conjugation, which is the definition of antiunitary symmetry (Porter 1965). One example is time reversal; some others are given by Robnik & Berry (1986).

Symmetrizability of A is a sufficient but not necessary condition for the absence of additional memory. One way in which the 2-form (11) might vanish for the nth mode, so that M is integrable for this mode, without A being symmetric or symmetrizable, would be if just one right eigenvector $\boldsymbol{u}_n$ were the same as its left counterpart, the other $N-1$ pairs being different.

If we abandon the assumption that the local eigenvalues q_n are non-degenerate, symmetrizability is no longer a sufficient condition. In the hermitian case, there can be additional memory for real symmetric matrices, in the form of a π phase shift (sign change) for circuits enclosing a degeneracy (Berry 1984). (In the non-hermitian case, degeneracies are branch points, and after a circuit the wave returns in a different mode; interesting memory effects (Berry 1990) can occur in this case also.)

4. Optically active refracting media: an instructive example

As their simplest example of a system claimed to exhibit additional memory, BS study a beam of light propagating in the z direction in an isotropic but inhomogeneous transparent medium with electric permittivity $\epsilon(z)$ and optical activity coefficient $\eta(z)$. The correct constitutive equation relating electric field and displacement in such a medium is

$$\begin{aligned}\boldsymbol{D} &= \epsilon_0\{\epsilon(z)\,\boldsymbol{E}+(1/2k)\,[\eta(z)\,\nabla\wedge\boldsymbol{E}+\nabla\wedge(\eta(z)\,\boldsymbol{E})]\}\\ &= \epsilon_0\{\epsilon(z)\,\boldsymbol{E}+(1/k)\,\eta(z)\,\nabla\wedge\boldsymbol{E}+(1/2k)\,[\nabla\eta(z)]\wedge\boldsymbol{E}\}.\end{aligned} \tag{21}$$

It is essential to include the term in $\nabla\eta$, in order that the constitutive relation be hermitian and thus consistent with energy conservation. The associated energy flow vector, for the complex field representing a monochromatic wave, is (Landau *et al.* 1984, ch. XII, especially p. 361)

$$\boldsymbol{P} = \mathrm{Re}\,[\boldsymbol{E}^*\wedge\boldsymbol{H}+\tfrac{1}{2}\mathrm{i}c\epsilon_0\,\eta(z)\,\boldsymbol{E}^*\wedge\boldsymbol{E}] \tag{22}$$

and it follows from Maxwell's equations that $\nabla\cdot\boldsymbol{P}=0$. Without the last term in (21), energy is not conserved and the constitutive relation cannot represent a transparent medium. BS omit this term and thus fall into error, as we shall see.

Maxwell's equations take the form (1) by choosing $\boldsymbol{e}$ as the 2-component vector

$$\boldsymbol{e} \equiv \begin{bmatrix}\surd\epsilon_0(E_x+\mathrm{i}E_y)\\ \surd\mu_0(H_x+\mathrm{i}H_y)\end{bmatrix} \tag{23}$$

and the matrix A as

$$\mathsf{A} = \begin{pmatrix}0 & \mathrm{i}\\ -\mathrm{i}\epsilon+\eta'/2k & \eta\end{pmatrix}. \tag{24}$$

This depends on the medium parameters $X=(\epsilon,\eta)$ and also contains the 'hermiticity' term involving η'. To the lowest non-vanishing order as $k\to\infty$, this term can be neglected when calculating the additional memory factor (9) in the field (4), but must be included, through the eigenvalues $q(z)$, in the ordinary phase memory.

The two eigenvalues of A are

$$q_\pm = \tfrac{1}{2}\eta\pm\surd(\epsilon+\tfrac{1}{4}\eta^2+\tfrac{1}{2}\mathrm{i}\eta'/k) = \tfrac{1}{2}\eta\pm\surd(\epsilon+\tfrac{1}{4}\eta^2)\pm\mathrm{i}\eta'/4k\,\surd(\epsilon+\tfrac{1}{4}\eta^2)+\dots. \tag{25}$$

Therefore the ordinary phase memory exponent in (4) is

$$-\mathrm{i}k\int_0^z \mathrm{d}z\, q_\pm(z) = -\mathrm{i}k\int_0^z \mathrm{d}z(\tfrac{1}{2}\eta \pm \surd(\epsilon+\tfrac{1}{4}\eta^2)) \pm \int_{\eta(0)}^{\eta(z)} \frac{\mathrm{d}\eta}{4\surd(\epsilon+\tfrac{1}{4}\eta^2)} + \ldots, \tag{26}$$

where the neglected terms vanish as $k \to \infty$. The last term, absent in BS, is a non-integrable modification to the amplitude (it has no i), and contributes to the ordinary memory through the requirement that the medium be transparent, which produces the extra term in $\boldsymbol{A}$.

According to (9), the additional memory factor M involves the eigenvectors of $\boldsymbol{A}$ and $\boldsymbol{A}^{\mathrm{T}}$, which from (24) are

$$\boldsymbol{u}_\pm = \begin{pmatrix} 1 \\ -\mathrm{i}q_\pm \end{pmatrix}, \quad \boldsymbol{v}_\pm = \begin{pmatrix} 1 \\ \mathrm{i}q_\pm/\epsilon \end{pmatrix}. \tag{27}$$

From (10) it now follows that

$$M = \exp\left\{-\int_{\text{path}} \frac{q_\pm\,\mathrm{d}q_\pm}{\epsilon+q_\pm^2}\right\} = \exp\left\{\mp\frac{1}{4}\int_{\text{path}}\left[\frac{\mathrm{d}\eta}{\surd(\epsilon+\tfrac{1}{4}\eta^2)} \pm \frac{\mathrm{d}\epsilon+\tfrac{1}{2}\eta\,\mathrm{d}\eta}{\epsilon+\tfrac{1}{4}\eta^2}\right]\right\}. \tag{28}$$

This is equivalent to the result obtained by BS, and by evaluating the 2-form (11) it can be shown to be non-integrable. But a non-integrable real additional memory is impossible in a transparent medium, because with repeated cycling it could be employed to progressively attenuate the wave, or, by reverse cycling, to amplify it; any additional memory would have to take the form of a geometric phase.

However, this additional memory must be combined with the last term of the ordinary memory (26), which gives

$$\begin{aligned} M\exp\left\{\pm\frac{1}{4}\int_{\text{path}}\left[\frac{\mathrm{d}\eta}{\surd(\epsilon+\tfrac{1}{4}\eta^2)}\right]\right\} &= \exp\left\{-\frac{1}{4}\int_{\text{path}}\frac{\mathrm{d}\epsilon+\tfrac{1}{2}\eta\,\mathrm{d}\eta}{\epsilon+\tfrac{1}{4}\eta^2}\right\} \\ &= \exp\left\{-\frac{1}{4}\int_{\text{path}}\mathrm{d}\ln g(\epsilon+\tfrac{1}{4}\eta^2)\right\} = \frac{(\epsilon+\tfrac{1}{4}\eta^2)_0^{\frac{1}{4}}}{(\epsilon+\tfrac{1}{4}\eta^2)_2^{\frac{1}{4}}}. \end{aligned} \tag{29}$$

We see that the non-integrability of M has been cancelled by that of the ordinary memory, and what remains is a purely local amplitude factor. The additional memory (28), without the cancellation, is thus revealed as the artefact of an unphysical constitutive equation. (Of course, it would have been possible – although less instructive – to eliminate both memory contributions from the beginning, by making an initial transformation to an 'adiabatic' basis.)

An isotropic medium with variable permittivity and optical activity is not rich enough to generate additional memory. However, as I have shown elsewhere (Berry 1986), if the medium is anisotropic, with an axis of birefringence, and an axis of gyrotropy that is locally fixed (e.g. by the direction of a magnetic field), then slow rotation of these axes will give rise to a geometric phase.

References

Berry, M. V. 1984 *Proc. R. Soc. Lond.* A **392**, 45–57.

Berry, M. V. 1986 In *Fundamental aspects of quantum theory* (ed. V. Gorini & A. Frigerio). NATO ASI Series vol. 144, pp. 267–278. New York: Plenum.

Berry, M. V. 1990 *Proc. R. Soc. Lond.* A **430**, 405–411.

Berry, M. V. 1991 *Phys. Today*. (In the press.)

Budden, K. G. & Smith, M. S. 1976 *Proc. R. Soc. Lond.* A **350**, 27–46. (Called BS in the text.)

Garrison, J. C. & Wright, E. M. 1988 *Phys. Lett.* A **128**, 177–181.

Landau, L. D., Liftshitz, E. M. & Pitaevskii, L. P. 1984 *Electrodynamics of continuous media*, 2nd edn. Oxford: Pergamon.

Markovski, B. & Vinitsky, S. I. 1989 *Topological phases in quantum theory*. Singapore: World Scientific.

Porter, C. E. 1965 *Statistical theories of spectra: fluctuations*. New York: Academic Press.

Robnik, M. & Berry, M. V. 1986 *J. Phys.* A **19**, 669–682.

Shapere, A. & Wilczek, F. (eds) 1989 *Geometric phases in physics*. Singapore: World Scientific.

Smith, M. S. 1975 *Proc. R. Soc. Lond.* A **346**, 59–79.

Received 27 September 1990; accepted 9 October 1990