E. WOLF, PROGRESS IN OPTICS XXXVII

VI

TUNNELING TIMES AND SUPERLUMINALITY

BY

RAYMOND Y. CHIAO

Department of Physics, University of California at Berkeley, Berkeley, CA 94720, USA

AND

AEPHRAIM M. STEINBERG

Department of Physics, University of Toronto, Toronto, ON M5S 1A7, Canada

CONTENTS

"Time, time, time, see what's become of me."
– P. Simon and A. Garfunkel

§ 1. Introduction

How long does it take for a particle to tunnel through a barrier? This simple-sounding question has provoked much controversy over the past six decades, ever since the phenomenon of tunneling (i.e., barrier penetration) was first predicted to occur in quantum mechanics. Although tunneling has by now been observed in many physical settings, and has even been applied in many useful devices – such as the Esaki tunnel diode, the Josephson junction, and the scanning tunneling microscope – the speed of the tunneling process remains controversial. One reason for this is that some theories for the tunneling time predict – and some experiments confirm – that the time is so short that (in a sense to be defined more precisely below) the tunneling process is superluminal.

The tunneling time question is not only of scientific, but also of technological interest. It is important to know if there is any limitation on the speed of electronic and photonic devices arising from the speed of the tunneling process. Although the tunneling of electrons seems to be more important at the present time for practical devices, the tunneling of photons is central to such devices as fiber couplers, laser output couplers, and scanning photon tunneling microscopes, to name a few examples in optics[1].

Many conflicting theoretical predictions have been made concerning the tunneling time, and as yet no consensus has emerged as to the theoretical answer. However, the situation is changing rapidly because many experiments, mainly in optics, have now been performed to measure the tunneling time, and the purely theoretical debate has been transformed into one in which actual data can be brought to bear on the question. In the process, it has become clear that one must make a careful *operational* definition of exactly *how* the measurement of the tunneling time is actually performed. To show that a clear operational definition is in fact possible at all, we give here one example (others may also be possible): Suppose that two particles were produced simultaneously from a radioactive decay. One particle tunnels through a barrier towards a detector, and

[1] We use the term "optics" throughout this chapter to refer to all electromagnetic propagation phenomena, including not only those in the visible but also in the microwave region of the spectrum.

the second particle propagates through the vacuum towards a second detector. If the two particles have the same speed, and the two detectors were placed at an equal distance from the radioactive source, the time delay between the two "clicks" registered in coincidence by these detectors would then constitute a clear operational measure of a tunneling time (Steinberg, Kwiat and Chiao [1993]).

However, different experimental setups may measure different tunneling times, and the answer to the tunneling time question may differ from experiment to experiment. In particular, one must distinguish carefully between a tunneling *arrival* time, which measures how long it takes a particle to cross the barrier and reach the detector, and a tunneling *interaction* time, which measures how long the tunneling particle interacts with the barrier while crossing it. While classical intuitions lead us to take for granted that these two times ought to be identical, there is no physical law which guarantees this. In fact, as we shall see, *Gedanken-experiments* designed to measure one or the other of these quantities will in general not agree in quantum mechanics. This is a subtle but important distinction.

Measurements of tunneling times by photons possess certain advantages over those by electrons or other particles, stemming mainly from the fact that the wavelength of visible light is much larger than the de Broglie wavelength of massive particles. (Only at temperatures on the order of microkelvins does the thermal de Broglie wavelength of even a light atom approach microns; see § 6.) This makes the relevant physical dimensions of the tunneling barrier much larger, and hence makes the barrier much easier to fabricate for photons than for electrons. The photon also possesses an internal degree of freedom, namely its spin, which could be used as an internal clock during the tunneling process. However, electrons possess certain advantages over photons, the main one being that they possess an electric charge, and therefore that they interact with other charged particles strongly, thus allowing an easier measurement of the tunneling interaction time.

There are three main types of tunnel barriers for photons which have been used in tunneling-time experiments: (i) periodic dielectric structures excited inside their band gap or stop-band, (ii) frustrated total internal reflection (FTIR) in glass or dielectric prisms, and (iii) waveguides beyond cutoff, which have been studied so far using microwaves only. The first type of barrier arises from Bragg reflections from the periodic structure, which leads to an evanescent (i.e., exponential) decay of the wave amplitude when the frequency is within the forbidden band gap (or "photonic band gap", Yablonovitch [1993], John [1991]) at the first Brillouin zone edge, analogous to the evanescent propagation of electron waves with energies inside the band gap of a Kronig–Penney model. It should be noted that within a large bandwidth near the midgap region, the periodic structure is nondispersive (i.e., the group velocity approaches a constant), so that tunneling wave packets which are tuned to midgap, although much attenuated in amplitude, can remain essentially undistorted upon transmission through the barrier. The second type of barrier (FTIR) arises from the coupling of an evanescent wave in the spatial gap between a pair of glass prisms when a beam of light is incident on the interface between the prisms beyond the critical angle (Zhu, Yu, Hawley and Roy [1986]). The third type arises from the evanescent wave inside a waveguide whose dimensions are too small to allow the propagation at the incident frequency. There is negligible dispersion of the tunneling wave packet in FTIR, but waveguides beyond cutoff are highly dispersive.

There are other situations besides the three mentioned above, in which tunneling-like phenomena occur in optics; for example, wave propagation below the plasma cutoff frequency, or into the shadow region of a sharp edge by diffraction, or outside the allowed orders of diffraction gratings or Fabry–Perots, or inside absorption lines. These are all "classically" forbidden phenomena, i.e., all are forbidden at the level of geometrical or ray optics, but all can actually occur at the level of physical or wave optics. Some of these phenomena have also been shown experimentally to be superluminal.

There has recently been a second controversy which has arisen as a result of those experiments in which superluminal group delays through tunnel barriers have been observed. This controversy is centered around a different, but related, question: Can one send signals, that is, information, through a tunnel barrier faster than the vacuum speed of light? This controversy has been sharpened by the claim by Nimtz and his co-workers that they have actually transmitted Mozart's 40th Symphony as a radio signal through a microwave tunnel barrier at a speed much faster than c (Heitmann and Nimtz [1994]). We shall show that there has in fact been no violation of Einstein causality in these and closely related experiments. Therefore the implication that their experiments somehow call causality into question is in our opinion unfounded.

In light of the second of these controversies, we have decided to include in this review a discussion of the problem of superluminal group velocities which have been predicted for the propagation of wave packets tuned to transparent spectral regions of media with inverted atomic populations. We shall discuss two cases: superluminal wave packets tuned close to zero frequency, and those tuned close to an atomic resonance with gain in it. In the latter case, optical tachyon-like propagation of collective atom-photon excitations is predicted to occur. These new kinds of superluminal propagation effects can occur over much longer distances than for tunneling. Hence they will force us to understand the meaning of causality, the definition of a signal, and the nature of information.

§ 2. A Brief History of Tunneling Times

Despite its absence from the overwhelming majority of textbook descriptions of tunneling (one early exception being Bohm [1951]), the tunneling-time problem has a long and illustrious history. At the heart of the problem is the fact that the kinetic energy of a particle inside a tunnel barrier is negative, so that a semiclassical estimate of its velocity becomes imaginary. This makes it impossible to make the naïve first approximation that the duration of a tunneling event is the barrier width divided by the velocity $\sqrt{2E/m}$.

Within a few years of the first predictions of tunneling, discussions appeared of the time spent by a particle in a "forbidden" region, and of the use of the stationary phase approximation to calculate properties of tunneling wave packets (Condon [1931], MacColl [1932]). By 1955, Wigner published a paper discussing the relationship between scattering phase shifts and the delay time, making explicit the connection between these quantities and the principle of causality (Wigner [1955]). As is well-known in electromagnetism, the frequency-derivative of the transmission phase shift yields the time-of-arrival of a well-behaved wave packet peak; we term this quantity the "group delay", by analogy with the group velocity calculated by the same method of stationary phase[2]:

$$\tau_g \equiv \frac{\partial \phi_t}{\partial \omega}, \tag{2.1}$$

where ϕ_t is the phase of the tunneling transmission amplitude. (For a free particle with a real momentum $\hbar k$, we have $\phi_t(x) = kx$, and the above relation yields $\tau_g(x) = x/(\mathrm{d}\omega/\mathrm{d}k)$, where the denominator is the familiar expression for the group velocity.) Wigner and his student Eisenbud [1948] applied the interpretation of the derivative of the scattering phase shift as a time delay to the problem of scattering (which of course includes tunneling as a special case), and Wigner observed that "the 'retardation'[3] cannot assume arbitrarily large negative values, in classical theory it could not be less than $-2a$", where a is the radius of the scattering potential; in other words, a classical particle cannot leave the scattering center before it arrives. Wigner noted that "It will be seen that the wave nature of the particles does permit some infringement of [this

[2] It is important to note, however, that many workers use the terminology "phase time". We avoid this, as the confusion between phase and group velocities has occasionally clouded the causality issues which plague the tunneling-time controversy.

[3] By 'retardation', Wigner refers to the group delay relative to the time for free propagation, expressed in units of distance.

inequality]". It is primarily with this very infringement that we are concerned here. Does wave mechanics truly allow particles to exit a barrier before they enter it, and in particular, do such effects violate relativistic causality? One could reasonably suspect that the non-relativistic nature of the Schrödinger equation is at fault here, but more careful analyses using the Dirac equation show that such superluminal transmission (which occurs in cases where all relevant energy scales are far less than the electron rest mass in the first place) persists (Leavens and Aers [1989]). The conflict is made even more clear by turning to optical analogs of tunneling, as the same problems arise with Maxwell's (fully relativistic) equations, and since one begins in the relativistic regime, it is relatively easy to achieve conditions under which the group delay is predicted to be superluminal.

Of course, superluminal and even negative group velocities were already known to occur in electromagnetism, and had been reconciled with causality by Sommerfeld and Brillouin (Brillouin [1960]). Their work showed that no real signal could propagate faster than the vacuum velocity of light c in any medium obeying the Kramers–Kronig relations, even in regions of anomalous dispersion. In these regions, the absorption and the strong frequency-dispersion cause the stationary-phase approximation to break down, as an incident pulse is distorted beyond recognition and no single transmitted peak may be observed. Conventional wisdom has it that such a breakdown occurs in every limit where the group velocity exceeds c. Nevertheless, as early as 1970, Garrett and McCumber showed theoretically that for short enough interaction lengths, absorbing media could indeed transmit undistorted (but attenuated) Gaussian pulses at superluminal, infinite, or even negative group velocities (Garrett and McCumber [1970]). The experimental verifications of these predictions will be discussed in § 4.2. As we shall see, these effects are consistent with relativistic causality, and no signal is in fact conveyed faster than light.

Even before this work on anomalous dispersion (which is still not as widely known as it deserves to be), the question of superluminal wave-packet transmission in tunneling was put on a firmer footing by Hartman [1962]. Hartman was not satisfied by MacColl's 1932 observation that there is "no appreciable" delay in tunneling[4], and he was concerned about the effects of preferential transmission of higher energy components in a wave packet. In a

[4] As alluded to earlier, it is the imaginary momentum which leads to difficulty. If k is imaginary, i.e., if the wave function decays exponentially according to $\Psi(x) = \Psi(0)e^{-\kappa x}$, the phase ϕ becomes a constant, and the group delay of eq. (1) vanishes, apart from effects due to boundary conditions.

rigorous treatment of the tunneling of wave packets through a rectangular barrier, he indeed found that for very thick barriers, such distortion occurred that no peak could be identified which might appear at the group delay time. For thin barriers, his results were in agreement with the stationary phase prediction, but there was no conflict with causality. Roughly speaking, the prediction is that for thicknesses smaller than one decay length of the evanescent wave ($d < 1/\kappa$), a transmitted particle of energy much less than the barrier height ($E \ll V_0$; $k \ll k_0$) will appear to have travelled at its initial velocity of $\hbar k/m$. This delay is related to the fact that phase is accumulated as the evanescent ($e^{-\kappa x}$) and anti-evanescent ($e^{+\kappa x}$) waves change in relative size, as the two have different (but constant) phases. For thicker (i.e., "opaque") barriers ($\kappa d \gg 1$), there is no phase change across most of the barrier, since the wave function is dominated by real exponential decay. The group delay thus saturates at the finite value $2m/\hbar k\kappa$, the time it would take the free incident particle to traverse two exponential decay lengths $1/\kappa$. Hartman confirmed that for intermediate barrier thicknesses, larger than $1/\kappa$ but small enough that the pulse was not distorted significantly, this saturation effect did indeed occur. As the distance traversed continues to grow, but the time required to traverse it remains roughly constant, it is clear that one eventually reaches a regime where the apparent propagation speed exceeds c. (Recently, Low and Mende [1991] argued that an actual measurement of such an anomalously short (tunneling) traversal time could not be made. However, Deutch and Low [1993] modified this conclusion in the case of relativistic particles.)

We now know that there are a number of cases in both electromagnetism and quantum mechanics where the naïve application of a causality limit to the description of a wave packet's propagation may fail. The question is no longer whether the method of stationary phase is valid, but rather whether it is unique. Are there perhaps a multiplicity of timescales which describe tunneling? Does the superluminal appearance of a wave packet peak imply an anomalously short "dwell time", or some other "interaction time", for the particle inside a forbidden region? What does it say about the transport of energy, or of information? What of the fundamental limit on the speed of a device whose operation depends on tunneling?

Over the years, and particularly since the 1980s, numerous proposals have been made for other "tunneling times" which might best describe the *duration* of the tunneling process, rather than just the time of appearance of a wave packet peak. Büttiker and Landauer, in particular, stressed that no physical law guarantees that an incoming peak turns into an outgoing peak (Büttiker and Landauer [1982]). They and other workers have argued strenuously that the group delay is not a physically significant timescale. This dispute becomes subjective, of course, as recent experiments have shown unequivocally that in measurements of arrival time, the group delay is indeed significant. Other experiments which involve tunneling in solid state physics seem to be best described by the Büttiker–Landauer or Larmor timescales (Guéret, Baratoff and Marclay [1987], Guéret, Marclay and Meier [1988a,b], Esteve, Martinis, Urbina, Turlot and Devoret [1989], Landauer [1989]). We are thus left in the uncomfortable situation of being unable to identify a unique timescale for tunneling, which forces us to analyze each conceivable experimental situation separately. The continued work on tunneling times is driven largely by the hope that this potentially infinite number of timescales can be reduced to a manageably finite handful of definitions, whose relationships and physical significances can be pinned down precisely. Although this project is by no means complete, recent work leaves us hopeful that this goal is not an unreasonable one, and that we will soon arrive at a fuller understanding of tunneling and related phenomena.

Most optical experiments on tunneling times have studied the group delay; in general, it is more straightforward to measure the arrival time of a photon or an electromagnetic wave than to measure the duration of its interaction with some barrier. In a complementary fashion, studies of tunneling in solid-state physics have so far been unable to observe the group delay, but have lent support to certain other proposed times. Here we focus primarily on the former, but we will also discuss to some extent other candidate times and the outlook for future experiments on them.

While it is impossible in this context to provide a full description of every theory that has appeared on the question of tunneling times, there are certain leading contenders with which it is useful to be familiar. The "dwell time" τ_{d} seems the most straightforward answer to the question "How much time does a particle spend in the barrier region?" It can be defined alternately for the time-dependent or the time-independent case. In the former, its natural statement is as the time-integral of the instantaneous probability that the particle is inside the barrier (assumed to extend from $-d/2$ to $d/2$):

$$\tau_{\mathrm{d}}\,(\text{time-dependent}) \equiv \int_{-\infty}^{\infty} \mathrm{d}t \int_{-d/2}^{d/2} \mathrm{d}x\,|\Psi(x,t)|^2 \,. \tag{2.2}$$

In the latter case, it is simply the probability density within the barrier, divided by the incident flux J_{in}:

$$\tau_{\mathrm{d}}\,(\text{time-independent}) \equiv \frac{1}{J_{\mathrm{in}}} \int_{-d/2}^{d/2} \mathrm{d}x\,|\Psi(x)|^2 \,. \tag{2.3}$$

In the limit of a monochromatic wave packet, these two formulas yield the same result, although for packets of finite extent, the corrections may be important (Hauge, Falck and Fjeldly [1987]).

The importance of definitions in the quantum regime cannot be overexaggerated. In the classical limit, τ_d (the time spent within the potential step) and τ_g (the time between arrival at the leading edge of the step and departure from the trailing edge) are of course identical, and equal to $d/v = md/\hbar k$. There is only one sensible quantity to term the "traversal time" in this case. It is because in the quantum limit these different definitions, equivalent in all familiar, classical regimes, yield different answers that there is no unambiguous recipe for providing an experimental prescription for determining "*the* tunneling time" quantum mechanically.

This difficulty has been traced most frequently to two characteristics of quantum mechanics. One is the fact that time is not an observable: there is no Hermitian operator corresponding to the time of arrival, or to the duration of an interaction. The other crucial characteristic is that unlike classical mechanics, quantum mechanics (or wave mechanics, more broadly) does not contain well-defined trajectories with determined durations. A particle's traversal of a barrier may be described as a Feynman path-integral (or Huygens'-Principle sum) over every possible trajectory linking its emission and its subsequent detection (Fertig [1990, 1993], Sokolovski and Baskin [1987], Sokolovski and Connor [1990, 1993], Hänggi [1993]). Since the different trajectories in general have different durations, we see that we should not necessarily hope to find a precisely defined interaction time for a quantum particle.

Nonetheless, we are free to consider specific experiments and ask by what timescales they are governed. In simple cases, we can perform the full quantum-mechanical analysis in order to arrive at a result. If we are fortunate, we may discern certain patterns in these results which will allow us to make inferences about problems too complicated for exact solution. At the least, by understanding some of these timescales, we hope to pin down the limits of validity of various approximations, such as the assumption that external degrees of freedom follow adiabatically the evolution of the tunneling particle, or in the opposite limit, remain unaffected by the motion of the particle.

The dwell time may appear unsatisfactory as a candidate for several reasons. Foremost, it is a characteristic of an entire wave function, comprising both transmitted and reflected portions. One might well expect that transmitted and reflected particles could spend differing amounts of time in the barrier. (Without a doubt, one would expect them to spend different amounts of time on the *far* side of the barrier – a finite amount for the transmitted particles and none for the reflected ones – whereas the formulations of eqs. (2.2) and (2.3) leave no room to introduce this distinction.) Its definition is so natural that many researchers have argued that it must at least reflect the *weighted average* of transmission and reflection times, $\tau_d = |t|^2\tau_t + |r|^2\tau_r$ (with t and r the transmission and reflection amplitudes, respectively), but even this assertion has been disputed hotly (Hauge and Støvneng [1989], Büttiker [1990], Sokolovski and Baskin [1987], Sokolovski and Connor [1990, 1993], Olkhovsky and Recami [1992], Landauer and Martin [1994]).

The second seeming problem with the dwell time is one it shares with the group delay. It is not guaranteed to be greater than the barrier thickness upon the speed of light, d/c. In fact, in the low-energy limit $k \to 0$, the wave is almost entirely reflected by the first interface, and $|\Psi|^2$ is negligible in the barrier, leading τ_d to vanish as $2mk/\hbar\kappa k_0^2$.

Büttiker and Landauer [1982] have been the great champions of looking beyond the group delay and the dwell time to definitions related more closely to the kinds of experimental questions which might concern us. In their 1982 paper, which is widely viewed as having rekindled the tunneling-time fire, they proposed a *Gedankenexperiment* which would allow one to infer the duration of the tunneling process. Consider a particle tunneling through a rectangular barrier. Now modulate the height of the barrier by a small amount, at some relatively low frequency Ω. Clearly, the transmission is lowest when the barrier is highest, and vice versa. But now imagine that Ω becomes greater and greater, until $\Omega \gg 1/\tau_t$, that is, until the barrier goes through more than one oscillation during the "duration" τ_t of the tunneling event. Naturally, the modulation of the transmitted wave will be washed out. Büttiker and Landauer therefore solved the problem of the oscillating barrier, and looked for this critical frequency Ω_c. They then postulated that the traversal time was $\tau_{BL} \equiv 1/\Omega_c$. When the calculation was performed in the opaque limit ($\kappa d \gg 1$), they found the following result:

$$\tau_{BL} = md/\hbar\kappa \,. \tag{2.4}$$

This is a striking result. Recalling that the local wavevector inside the barrier is $i\kappa$, we see that this is exactly the time we would expect from a semiclassical or WKB approach ($md/\sqrt{2mE}$) – aside from the fact that we find a real number here, despite the imaginary value of the wavevector. Due to the similarity of the formulas, the Büttiker–Landauer time is also frequently referred to as the "semiclassical time". (Far above the barrier, both τ_g and τ_d in fact approach the semiclassical time $\tau_s \equiv md/\hbar|k|$.) Since this time is proportional to d, it rarely becomes smaller than d/c; in fact, it would only do so for $m/\hbar\kappa > c$, which is the relativistic limit, where the Schrödinger equation should not be expected to be valid. (In reality, for geometries more complicated than the rectangular barrier, it has been noted that this time may vanish identically, leading once more to causality problems (Büttiker and Landauer [1985], Støvneng and Hauge [1989], Martin and Landauer [1992], Steinberg, Kwiat and Chiao [1993]).) While above the barrier, the semiclassical time closely resembles the group delay (missing only the oscillations due to multiple reflections at the barrier edges, which become insignificant in the WKB limit), it looks nothing at all like τ_g below the barrier, diverging when $E = V_0$ (where $V_0 \equiv \hbar^2 k_0^2/2m$ is the height of the barrier) and falling in the opaque limit ($\kappa d \gg 1$) to $md/\hbar k_0$ as opposed to diverging like $\tau_g \rightarrow 2m/\hbar k k_0$. The group delay diverges for $k \rightarrow 0$, but is independent of d; the Büttiker–Landauer time is well-behaved as $k \rightarrow 0$, and is proportional to d.

Büttiker went on to consider another "clock", to see if different types of perturbations would bring to light the same timescale. Expanding on work due to Baz' [1967] and Rybachenko [1967], he considered an electron tunneling through a barrier to which a small magnetic field $\boldsymbol{B} = B_0\hat{\boldsymbol{z}}$ is confined. Suppose the electron's spin is initially pointing along $\hat{\boldsymbol{x}}$. The magnetic field causes it to precess in the x–y plane at the Larmor frequency $\omega_L = 2\mu_B B_0/\hbar$, where μ_B is the Bohr magneton. If one measures the polarization of the transmitted electron, one will find it to have precessed through some angle θ_y, and nothing could be more natural than to ascribe this to precession at ω_L for the duration τ_y of the tunneling event, leading to the "Larmor time" $\tau_y \equiv \theta_y/\omega_L$. This time turned out to be equal to the dwell time τ_d, including the latter's superluminal behavior at low energies. (For cases other than the simple rectangular barrier, these two times do not remain equal. Hence some workers (Hauge and Støvneng [1989]) have argued that they are conceptually quite distinct quantities.)

Büttiker's insight was that this early expression for the Larmor time made the implicit assumption that by taking the $B_0 \rightarrow 0$ limit, one could neglect the tendency of the electron to align itself with respect to the magnetic field. In reality, due to the interaction Hamiltonian $\mathcal{H}_{int} = +2\mu_B B_0 S_z$, a spin-up electron sees an effective potential with a higher barrier than that seen by a spin-down electron, and therefore has a lower transmission probability. As the $\hat{\boldsymbol{x}}$-polarized electrons are equal superpositions of $S_z = \pm 1/2$, this preferential transmission will tend to rotate the polarization out of the x–y plane towards the negative z-axis, so that the transmitted electron beam is slightly spin-polarized antiparallel to the applied $\boldsymbol{B}$ field. Büttiker showed that both this out-of-plane rotation and the in-plane precession were first-order in B_0, and furthermore, that the former dominated the latter in the opaque limit. Defining a second Larmor time related to the polar rotation according to $\tau_z = \theta_z/\omega_L$, he found this timescale to reproduce the $md/\hbar\kappa$ behavior he and Landauer had already calculated by considering the modulated barrier. Suggesting that the true interaction time should take into account the full three-dimensional rotation of the electron's spin, he proposed that the interaction time was $\tau_x \equiv \sqrt{\tau_y^2 + \tau_z^2}$. We refer to this time as "Büttiker's Larmor time" $\tau_{\rm B}$. It agrees with the oscillating-barrier result $\tau_{\rm BL}$ in both the low- and high-energy limits.

A fair number of other approaches had been tried by 1990, mostly yielding combinations of the timescales already described: the group delay, the dwell time, the in-plane Larmor time, the Büttiker–Landauer (or semiclassical) time, or Büttiker's Larmor time. For example, a Feynman-path approach in which the duration of all relevant paths was averaged with the weighting factor $\exp\{iS[x(t)]/\hbar\}$ yielded the "complex time" $\tau_c = \tau_y - i\tau_z$ (Sokolovski and Baskin [1987], Sokolovski and Connor [1990, 1993], Fertig [1990, 1993], Hänggi [1993], Sokolovski [1995]). It is easy to observe that the magnitude of this time is Büttiker's Larmor time, while its real and imaginary parts are (for rectangular barriers) the dwell time and minus the semiclassical time, respectively. (An earlier approach (Pollak and Miller [1984], Falck and Hauge [1988]) yielded a similar complex time, whose real part was the group delay, rather than the dwell time.) Despite this telling relationship, many found the concept of a complex time to be unphysical and rejected it out of hand. The similarity of such different approaches can be traced to a particularly convenient functional form (Büttiker [1983], Landauer and Martin [1994]) in which they can be written:

$$\tau_{\rm g} = \hbar\frac{\partial}{\partial E}\arg(t), \tag{2.5}$$

$$\tau_y = -\hbar\frac{\partial}{\partial V_0}\arg(t) = \tau_{\rm d} \rightarrow \tau_{\rm g} = \tau_s \text{ in WKB limit}, \tag{2.6}$$

$$\tau_x = -\hbar\frac{\partial}{\partial V_0}\ln|t| \rightarrow \tau_{\rm BL} \text{ in opaque limit}, \tag{2.7}$$

$$\tau_c = i\hbar\frac{\partial}{\partial V_0}\ln t, \tag{2.8}$$

$$= \tau_y - i\tau_z\,. \tag{2.9}$$

The group delay is the derivative of the transmission phase with respect to the particle's energy, while the in-plane Larmor time is the derivative with respect to the barrier height. Since the out-of-plane Larmor precession arises from preferential transmission of anti-aligned rather than aligned spin components, it can be expressed similarly as a frequency-derivative of the transmission *probability*.

This reflects the theoretical situation when optical experiments on tunneling began returning results around 1990. More recent work has begun to shed some light on why the different times are related in the way they are, and how one might physically interpret the real and imaginary parts of a complex time. This approach, and possible experiments, will be discussed in § 5 and § 6. At least one other principal theoretical approach deserves mention, and this one is sufficiently distinct that we have saved it for the end. It is clear that in classical mechanics, a particle follows a well-defined trajectory, and that such a trajectory can be defined as a certain approximation to the motion of a quantum-mechanical wave packet in the classical limit. The breakdown of such a notion leads to the difficulties regarding the quantum-mechanical tunneling time, in particular to the fact that a time defined in terms of wave packet arrival no longer need coincide with a time defined in terms of a clock which evolves while the particle is within the barrier. The most familiar treatment of trajectories in quantum mechanics is the Feynman path integral discussed above, according to which a particle follows *every possible* trajectory with a given weighting[5]. There is nevertheless a very different proposal for incorporating trajectories into quantum mechanics. This is the pilot wave model of Bohm and de Broglie (Bohm [1952], Holland [1993]). This deterministic interpretation of quantum mechanics invokes a dual reality, consisting both of the wave function Ψ (determined in the usual manner) and of a particle with a perfectly well-determined position. An ensemble of particles with initial positions described by the probability distribution $P(x,0) = |\Psi(x,0)|^2$ evolves deterministically according to the hydrodynamic equation of motion

$$\dot{x}(t) = -\mathrm{i}\frac{\hbar}{m}\nabla\Psi(x,t)\,, \tag{2.10}$$

which is sufficient to ensure that at all later times, the Born interpretation of $|\Psi|^2$ will remain valid. Although the ensemble as a whole is described by a wave function, and does not possess a unique traversal time, each individual particle follows a classical trajectory whose duration may be calculated. This approach has been followed by various workers (Dewdney and Hiley [1982], Leavens [1990], Leavens and Aers [1993]), and has been shown to have interesting

[5] Note that this weighting affords a rigorous prescription for calculating transition amplitudes, but no established recipe existed for defining a "duration", in the absence of a clear operational definition of the latter. It is a pleasant surprise then that the natural extension proposed by Sokolovski and Baskin [1987], Sokolovski and Connor [1990, 1993] and Hänggi [1993] agrees at any rate with other more or less justifiable definitions.

relationships to the times already discussed. In general, however, the physical significance of these trajectories remains an issue of some contention (see, for example, Englert, Scully, Sussmann and Walther [1992], Dürr, Fusseder, Goldstein and Zanghi [1993], Dewdney, Hardy and Squires [1993], Steinberg, Kwiat and Chiao [1994]). The one feature of the Bohm approach which makes it somewhat haunting is that different Bohmian particles from the same ensemble may not cross each other's trajectories, thanks to the single-valued velocity function given above. This implies that all transmitted particles originate earlier in the wave packet than all reflected particles. Given the superluminal behavior of tunneling peaks, it is striking that the particles which form the transmitted peak do *not*, under this interpretation, originate in the incident peak, but rather earlier in time. Later on, we shall see a similar feature in the classical-wave (pulse-reshaping) description of tunneling.

§ 3. Tunneling and Its Optical Analogs

We establish here on a more formal basis the analogy between electron and photon tunneling (Chiao, Kwiat and Steinberg [1991]). From Maxwell's equations for classical electromagnetic fields, one can derive the wave equation in an inhomogeneous but isotropic medium, which for a monochromatic wave in the scalar approximation reduces to the Helmholtz equation,

$$\nabla^2 \mathcal{E} + \{n(x,y,z)^2 \omega^2/c^2\}\mathcal{E} = 0 , \tag{3.1}$$

where $\mathcal{E}$ is the scalar amplitude of the electric field, $n(x,y,z)$ is the index of refraction of the medium at ω, the angular frequency of the wave, and c is the speed of light in the vacuum. The coefficient of $\mathcal{E}$ in the second term (the curly brackets) represents the square of the local wavevector. This equation is formally identical to the time-independent Schrödinger equation for the electron,

$$\nabla^2 \Psi + (2m/\hbar^2)\{E - V(x,y,z)\}\Psi = 0 , \tag{3.2}$$

where Ψ is the wavefunction of the electron, m is its mass, $V(x,y,z)$ is the potential energy, and E is the total energy. This identification is exact if we make the following identification[6]:

$$n(x,y,z) \Leftrightarrow \{2m[E - V(x,y,z)]\}^{1/2} c/\hbar\omega . \tag{3.3}$$

[6] Note, however, that the correspondence depends explicitly on ω, and thus is only exact over restricted bandwidths. A dielectric interface may have a reflectivity which tends to a constant less than one as the photon energy vanishes, while a step potential will always have reflectivity tending to unity as the electron energy vanishes. It can therefore be subtle to connect Kramers–Kronig-style arguments for photons to those for electrons.

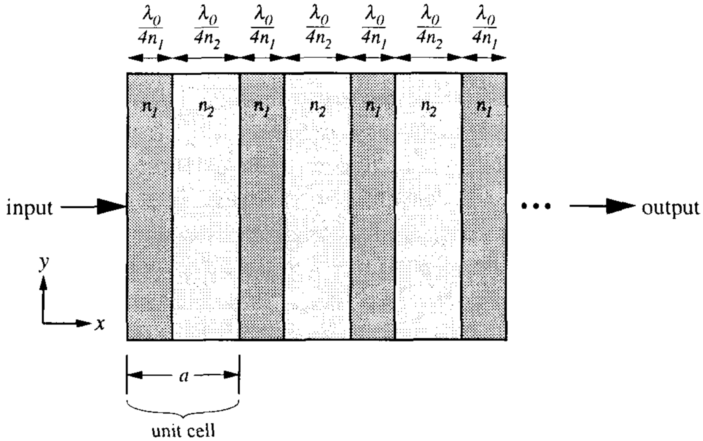

Fig. 1. Periodic stack of quarter-wave dielectric layers composed of alternating high- and low-index media, i.e., the 1D photonic band-gap material.

Tunneling barriers can arise in regions of space where $E < V(x,y,z)$, which correspond to evanescent wave regions, where the effective index of refraction $n(x,y,z)$ is imaginary. Several situations in optics give rise to such evanescent waves, and hence to photon tunneling. All involve propagation of waves beyond some sort of cutoff, such as the cutoff at a photonic band gap edge, the cutoff at the critical angle for total internal reflection, or the cutoff of a constricted waveguide.

As our first example of an optical tunneling barrier, we consider the evanescent wave propagation of electromagnetic waves inside a 1D photonic band gap, since there is an obvious analogy to the evanescent propagation of electrons inside the band gap of the Kronig–Penney model for periodic electronic structures. Let the photonic band-gap material be composed of two media with $n_1 > n_2$, described by

$$\begin{aligned} n(x,y,z) &= n_1 \quad \text{for all } ma \le x < ma + \frac{\lambda_0}{4n_1}, \\ n(x,y,z) &= n_2 \quad \text{for all } ma + \frac{\lambda_0}{4n_1} \le x < ma + \frac{\lambda_0}{4n_1} + \frac{\lambda_0}{4n_2}, \end{aligned} \tag{3.4}$$

where $m = 0, 1, 2, \ldots$, where λ_0 is the vacuum wavelength, and where the lattice constant a of the unit cell is given by

$$a = [n_1^{-1} + n_2^{-1}]\lambda_0/4. \tag{3.5}$$

This periodic dielectric stack is illustrated by fig. 1, and is equivalent to a dielectric mirror consisting of a periodic stack of alternating high- and low-index

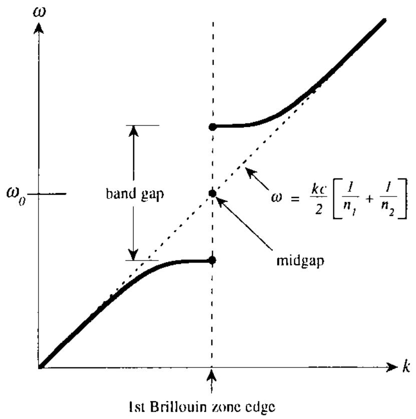

Fig. 2. Dispersion relation for the 1D photonic band-gap material, where the midgap frequency is $\omega_0 = 2\pi c/\lambda_0$.

quarter-wave layers. By eq. (3.3), we see that this is equivalent to the problem of an electron in a periodic potential, which can be approximated by the Kronig–Penney model (Ashcroft and Mermin [1976]). There results a band gap at the first Brillouin zone edge (see fig. 2) which arises from Bragg reflections off the periodic planes between the index strata. Hence the propagation of light inside the band gap becomes evanescent.

As a second example, we consider the case of frustrated total internal reflection (FTIR). Consider two right-angle glass prisms, which are placed with their hypotenuses in close proximity, so that coupling through the exponential tail of the light wave (for incidence angles beyond the critical angle) allows the leakage of light from one glass prism into the other through an air gap (see fig. 3). This case is easier to connect with textbook descriptions of tunneling, and has also been used in a number of recent experiments on tunneling times.

In the case of TE- or s-polarized light incident in the x–y plane on a glass–air interface at an angle θ (see fig. 3), we can take out the dependence of the electric field $\mathcal{E}\hat{z}$ on time and on y (the direction parallel to the interface) as follows:

$$\mathcal{E}(x,y,t) = \Psi(x)\,\mathrm{e}^{\mathrm{i}(ky\sin\theta-\omega t)} \tag{3.6}$$

in all three regions, where $k = n\omega/c$ is the wavevector in the glass. For s-polarization, where the electric field vector is perpendicular to the plane of incidence, $\mathcal{E}$ and thus Ψ are continuous across the boundaries. If we assume a magnetic permeability of $\mu = 1$ in all three regions, then the magnetic field $\mathcal{B}$ is continuous as well, and this leads to the continuity of $\Psi'(x) \equiv \mathrm{d}\Psi/\mathrm{d}x$.

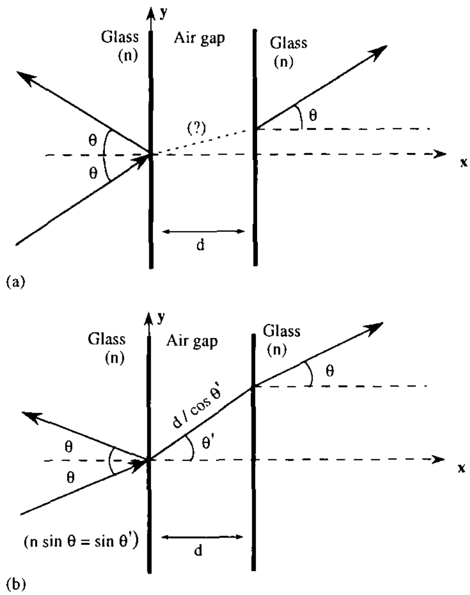

Fig. 3. Glass–air–glass interface with light rays drawn for the case of (a) tunneling through the air gap in frustrated total internal reflection (FTIR) when $\theta > \theta_c$, and (b) "classically allowed" transmission when $\theta < \theta_c$.

These boundary conditions are the same as those for the one-dimensional Schrödinger wave function $\Psi(x)$ at a step discontinuity in the potential $V(x)$. The electromagnetic wave equation reduces to

$$\begin{aligned} \Psi'' + (\omega/c)^2\{n^2\cos^2\theta\}\Psi &= 0 \text{ in glass regions} \\ \Psi'' + (\omega/c)^2\{1 - n^2\sin^2\theta\}\Psi &= 0 \text{ in the air gap}, \end{aligned} \tag{3.7}$$

where the coefficients of Ψ in the second terms represent the squares of the x-components of the wavevectors in the glass and in the air gap, respectively. Equation (3.7) has exactly the form of the one-dimensional Schrödinger equation for an electron in a rectangular barrier of height V_0 and a width equal to the width of the air gap (see fig. 3), when we draw the equivalences

$$\begin{aligned} 2mE/\hbar^2 &\Leftrightarrow (\omega/c)^2\{n^2\cos^2\theta\} \\ 2m(E - V_0)/\hbar^2 &\Leftrightarrow (\omega/c)^2\{1 - n^2\sin^2\theta\}. \end{aligned} \tag{3.8}$$

It is clear from this correspondence that the critical behavior at $E = V_0$ is analogous to that at the critical angle $\theta = \theta_c \equiv \sin^{-1}(1/n)$, and that for given

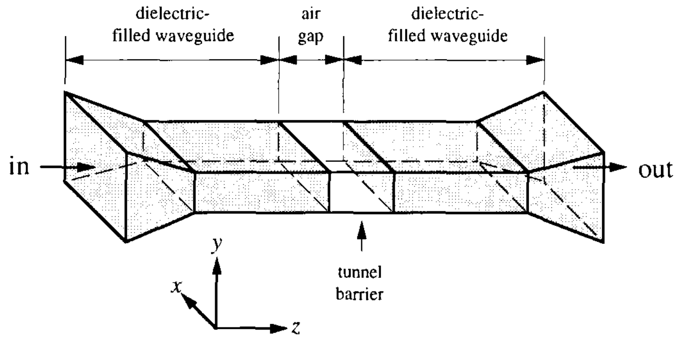

Fig. 4. Microwave tunnel barrier consisting of an air gap section between between two dielectric-filled sections of a rectangular wave guide.

electron mass and photon frequency, a precise one-to-one mapping can be made between the parameters E and V_0 of the electron experiment and the parameters θ and n of the photon experiment. In addition, in the classically allowed regime $E > V_0$, the velocity of the electron inside the barrier is proportional to $(E - V_0)^{1/2}$ in the classical (i.e., WKB) limit. When eq. (3.8) is used to transform this into the analogous photon variables, this electron velocity is seen to be proportional to $\cos\theta'$, where θ' is the angle of the refracted beam of light inside the air gap in the "classical" (i.e., geometrical optics) limit for the photon (see fig. 3b). Thus the electron traversal time mimics exactly this "ray optics" behavior of the corresponding photon traversal time (Steinberg and Chiao [1994a]). This is true in spite of the fact that their dispersion relations $E(p)$ are quite different.

As a third example, we consider a wave guide beyond cutoff. In order to avoid the complications of the fringing fields associated with a sudden decrease and increase in wave guide width, which is usually utilized in microwave experiments on the tunneling time, we analyze here instead the simpler case introduced by Martin and Landauer [1992], who considered a dielectric-filled wave guide interrupted by a rectangular air gap which serves as the barrier (see fig. 4). For simplicity, consider the TE_{10} mode of this wave guide. The dispersion relations come from the relationships

$$\begin{aligned} k_x^2 + k_z^2 &= n^2\omega^2/c^2 \quad \text{for the dielectric-filled sections,} \\ k_x^2 + k_z'^2 &= \omega^2/c^2 \quad \text{for the air gap,} \end{aligned} \tag{3.9}$$

where n is the index of refraction associated with the dielectric, and where the conducting boundary conditions impose the condition $k_x = \pi/a$ (a being the

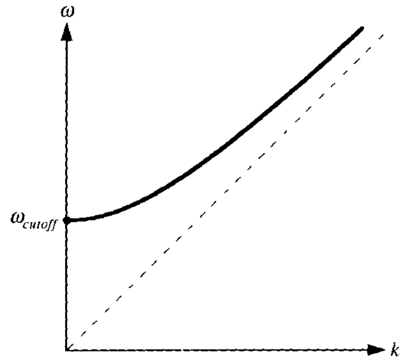

Fig. 5. Dispersion relation for the TE_{10} mode of the rectangular wave guide.

width of the wave guide) for the TE_{10} mode. Therefore the dispersion relation of the wave guide in the air gap is of the form (see fig. 5)

$$\omega^2 = \omega_c{}^2 + k_z'^2 c^2, \tag{3.10}$$

where $\omega_c = \pi c/a$. If the frequency of the wave is chosen to be below this cutoff, but above the cutoff frequency of the dielectric-filled section, then k_z' will be imaginary, while k_z is real, and this wave guide configuration becomes a good analog for the tunneling of an electron through a one-dimensional rectangular barrier. The group delay for this wave guide geometry has been calculated by Martin and Landauer [1992].

§ 4. Optical Experiments on Tunneling Times

4.1. CARNIGLIA AND MANDEL'S FTIR EXPERIMENT

An early optical experiment measuring the phase shifts which occur in frustrated total internal reflection (FTIR) was performed for both the TM and TE polarizations of the incident light (Carniglia and Mandel [1971a,b]). In a theoretical analysis of their experiment, Carniglia and Mandel calculated the time of arrival of the *phase* front of the evanescent wave at a point straight across the gap at a minimum distance from the point of incidence. Although their work did not directly address the problem of tunneling times, their results did bear indirectly on the question of whether or not the group delay saturates with increasing barrier thickness.

Using a modified Rayleigh interferometer, they measured the phase shift accumulated by an evanescent electromagnetic wave after it crosses the air gap between the two glass prisms. Because the evanescent wave propagates parallel to the glass-to-air interface, this wave can penetrate into a direction normal to the interface without much change of phase, since the dominant exponential decay of the evanescent wave amplitude is a real function. This was confirmed in their first experiment, in which they showed that for TM polarization the phase shift saturated at the theoretically predicted (asymptotic) value of

$$\phi_{\text{sat}} = \tan^{-1}\left[\frac{\cos^2\theta - n^2(n^2\sin^2\theta - 1)}{2n\cos\theta(n^2\sin^2\theta - 1)^{1/2}}\right], \tag{4.1}$$

which is independent of the width of the gap (i.e., the barrier thickness), in the opaque or thick-barrier limit.

Since the derivative of the phase with respect to the frequency is the group delay, their observation implied that the group delay should also saturate, and thus become independent of the barrier width. Thus their experimental result was consistent with the theoretical conclusion reached earlier by Hartman [1962]. Since there should be a crossover point beyond which the saturated group delay becomes less than the light-transit time across the barrier, these early experimental and theoretical papers already implied that the tunneling group delay should become superluminal for sufficiently thick barriers. In fact, since eq. (4.1) is independent of frequency, the saturated group delay is approximately zero. This implied that superluminal group delays should be easily achievable.

There is an additional contribution to the group delay arising from a lateral shift of the beam due to the Goos–Hänchen shift (Steinberg and Chiao [1994a]). This shift has been observed recently in the transmitted beam in FTIR by Balcou and Dutriaux (see § 4.11), and used by them to measure one of their two tunneling times. However, in Carniglia and Mandel's original experiment, the beam width was 6 cm, which was so large that they could not observe this shift.

4.2. ABSORPTIVE MEDIA WITH ANOMALOUS DISPERSION

In another optical context, superluminal group delays were also predicted theoretically and observed experimentally, namely in the region of anomalous dispersion near the center of an absorption line. Although this is not related directly to the question of tunneling times, many aspects of this earlier controversy concerning superluminal group velocities reappear in the tunneling-time controversies. In 1970, Garrett and McCumber returned to an old problem first considered by Sommerfeld [1907]. They showed theoretically that for short lengths, absorbing media could transmit undistorted (but attenuated) Gaussian pulses at superluminal, infinite, or even negative group velocities (Garrett and McCumber [1970]). This arose from the fact that the group velocity, which is given by the expression

$$v_{\mathrm{g}}(\omega) = c\left[\mathrm{Re}\, n(\omega) + \omega \frac{\mathrm{d}\, \mathrm{Re}\, n}{\mathrm{d}\omega}\right]^{-1}, \tag{4.2}$$

can have a vanishing denominator in regions of anomalous dispersion, where $\mathrm{d}\,\mathrm{Re}\, n/\mathrm{d}\omega$ is large and negative, i.e., near the center of a strong absorption line.

The stationary phase approximation does not automatically break down for smooth Gaussian pulses, in contrast to the case of signals with a discontinuous front considered by Sommerfeld and Brillouin (Brillouin [1960]). Garrett and McCumber showed that an incident Gaussian wave packet can be reshaped by the absorption process (in which the later parts of the wave packet would be absorbed to a greater extent than the earlier parts) in just such a way as to produce a smaller, but undistorted Gaussian wave packet at the exit face of the medium. (In tunneling, a similar pulse-reshaping occurs, except that the process of absorption is replaced by the process of attenuation due to reflection from the barrier.) The peak of the pulse thus appears to have moved at a superluminal group velocity inside the medium (or a barrier). Tanaka [1989] later extended their work using the saddle point method. He showed that the propagation of a wave packet into an anomalous dispersion medium is characterized by three successive spatial regions with negative, superluminal, and subluminal group velocities, respectively.

Chu and Wong [1982] verified experimentally that the superluminal behavior of the group velocity as predicted by these theories actually occurred for weak picosecond laser pulses propagating near the center of the bound A-exciton line of a GaP:N sample. Segard and Macke [1985] also confirmed these predictions in the propagation of millimeter wave pulses through a gas cell of OCS near the 97 GHz $J = 7 \rightarrow 8$ transition. Furthermore, both groups observed *negative* group velocities. The meaning of a negative group velocity is that the peak of the transmitted wave packet leaves the exit face of the gas cell *before* the peak of the incident wave packet enters the entrance face of this cell, in seeming defiance of our usual notions of causality. However, this effect can again be understood in terms of a pulse reshaping of the Gaussian wave packets due to absorption, and is perfectly causal (see § 8). These experiments demonstrated that the group velocity, even when it exceeds c, approaches infinity, or becomes negative, possesses a definite physical meaning, since there exist definite *operational* procedures, which have in fact been carried out in practice, to measure these counterintuitive group velocities. These facts fly in the face of conventional wisdom[7], which tells us that when the group velocity becomes superluminal, it has no longer any appreciable physical significance, or that somehow it is just not a useful concept.

4.3. THE MILWAUKEE GROUP

Starting in 1989, a group at Marquette University in Milwaukee began to generate a fair amount of controversy by publishing papers with titles as provocative as "Transmit radio messages faster than light". Needless to say, these articles were greeted with a great deal of skepticism, not mitigated by the fact that they seemed to harbor a confusion between phase and group velocities (Giakos and Ishii [1991a–c], Ishii and Giakos [1991], Stephan [1993]). Most physicists remained blissfully unaware of the argument, which nevertheless raged for a time in *Microwave and Guided Wave Letters*. The claims were twofold. The authors pointed out that for an electromagnetic wave propagating in free space, the phase velocity measured at an angle θ to the propagation direction is $c/\cos\theta > c$. They then claimed to have measured the arrival time for a microwave pulse in this geometry, and found it to be described by this superluminal phase velocity. They also did an experiment in a waveguide, presenting similar conclusions. Although they made no attempt to connect these findings to the phenomenon of tunneling, and their claims were not widely accepted, it is interesting to note that under certain conditions, such setups can indeed be shown to be analogous to tunneling, and to be described by time delays which in the appropriate limits become superluminal.

4.4. THE FLORENCE GROUP, PART I

Similar experiments were being carried out in a different spirit at the Istituto di Ricerca sulle Onde Elettromagnetiche del Consiglio Nazionale delle Ricerche in Florence at about the same time (Ranfagni, Mugnai, Fabeni and Pazzi [1991]). Ranfagni and co-workers were looking specifically at microwave transmission in waveguides beyond cutoff, whose mathematical equivalence to quantum-mechanical tunneling has already been noted. Aware of the controversy over

[7] See for example p. 23 of Born and Wolf [1975], or p. 302 of Jackson [1975].

tunneling, they hoped to resolve the issue by measuring the transmission delay time and comparing it to the group delay, the semiclassical or Büttiker–Landauer time, and Büttiker's Larmor time. Their initial results were for an abrupt step being transmitted through a 10-cm-long waveguide with a cutoff of 9.494 GHz, as much as 43 MHz above the incident frequency. Complicated by the abrupt (roughly 5 ns) turn-on of their step and by the dissipation in the waveguide, their results were inconclusive, but showed rough agreement with the semiclassical time. Theoretical work taking dissipation into account (Ranfagni, Mugnai, Fabeni and Pazzi [1990], Mugnai, Ranfagni, Ruggeri and Agresti [1992]) yielded reasonable agreement with the experimental data.

Refinements of this experiment (Ranfagni, Mugnai, Fabeni, Pazzi, Naletto and Sozzi [1991]) improved the signal-to-noise ratio, allowing good data to be obtained as far as 100 MHz below the cutoff of a 15-cm narrowed waveguide segment. These data clearly contradicted the divergent behavior of the semiclassical time at cutoff, and seemed to agree better with the group delay than with the other candidate times. The barrier was not thick enough, however, for the contradiction between the group delay theory and the naïve application of the causality principle to be checked directly. The Florence group also indirectly studied τ_z, the out-of-plane portion of the Larmor time (equivalent to the imaginary part of the complex times discussed earlier), and were able to confirm that it behaved as predicted as well. Their conclusions were therefore appropriately cautious: "... there is agreement between the experiments and the appropriate theoretical models. This fact ... leaves the identification of the tunneling time ambiguous". Furthermore, in this series of experiments, it was impossible to directly test the question of superluminality.

4.5. THE COLOGNE GROUP, PART I

While Ranfagni's group was working to extend their step-function transmission-time measurements further below cutoff to adjudicate between the semiclassical and group-delay theories, a group in Cologne was also using microwaves to study tunneling, aiming in particular to test the prediction of superluminal traversal. In their initial experiments (Enders and Nimtz [1992]), they used a network analyzer to measure the transmission phase shift through a narrowed waveguide at different frequencies. They inferred a group delay by fitting their phase data to a smooth curve, and subsequently performing a Fourier transform to predict the delay for a hypothetical pulse. For the longest barrier they used, 10 cm, they calculated a group delay of 130 ps, which would correspond to transmission at about 2.5 times the speed of light. They also observed, in agreement with the saturation effect predicted by Hartman, that barriers of different lengths yielded essentially the same phase shifts. In their early work, technical considerations made direct time-measurements less reliable than the phase measurements. In 1993, however, they reported time-domain measurements confirming the frequency-domain results, under the slightly misleading title "Zero-time tunneling of evanescent mode packets" (Enders and Nimtz [1993]). In this experiment, they used a Hewlett-Packard synthesizer to produce sharp-onset pulses (rise times of a few nanoseconds) with carrier frequencies near 8.65 GHz, allowed the waves to tunnel through a 6-cm barrier formed by a waveguide section with a 9.49 GHz cutoff frequency (with an attenuation of 40 dB), and then used a Hewlett-Packard transition analyzer to detect the transmitted envelope and compare it with that of a wave which traversed a 40 dB attenuator (whose effect on the group delay was verified independently to be negligible), but no barrier region. Due to the large bandwidth of their pulses, they saw a fair amount of distortion, and complicated features, but over much of the step, they found a propagation delay which appeared to be small relative to the 0.2-ns free-space propagation delay. They took this as final confirmation that the microwaves traversed the narrowed waveguide superluminally (indeed, with zero delay, since in a sense all the residual group delay may be attributed to edge effects, i.e., impedance mismatch between the waveguide segments).

4.6. THE BERKELEY GROUP

While most work on optical tunneling was going on with classical electromagnetic waves, typically in the 10 GHz range, at Berkeley we had proposed to perform a test of optical tunneling that would stress the single-particle aspects of the effect. Quantum electrodynamics predicts that for purely linear optical effects, such as those considered in this chapter, single photons exhibit the same behavior as classical pulses (Glauber [1965]). In fact, one may consider the (properly normalized) pulse profile as the single-photon wave packet[8]. It is possible to construct creation and annihilation operators for any pulse mode which is a solution of Maxwell's equations, simply by superposing operators

[8] Although the existence of "wave packets" for photons is controversial, it is possible by limiting oneself to cases where photon number is conserved and to the paraxial limit to consider the positive-frequency part of the electric field $E^+(\boldsymbol{r}, t)$ analogous to a quantum wave function, bearing in mind that the detection probability is proportional to $E^-(\boldsymbol{r}, t)E^+(\boldsymbol{r}, t) = |E^+(\boldsymbol{r}, t)|^2$, in analogy with the standard Born interpretation of the electron wave function (Deutsch [1991], Deutsch and Garrison [1991]).

for the plane-wave modes (Deutsch [1991], Deutsch and Garrison [1991]). Propagation effects are then governed by the classical wave equations, and quantization merely affects detection *statistics* and higher-order effects. Having already shown (Steinberg, Kwiat and Chiao [1992a]) that single-photon wave packets travelled at the group velocity in media with normal dispersion, we decided to extend this work to the case of tunneling.

Our original proposal (Chiao, Kwiat and Steinberg [1991]) discussed the analogy between frustrated total internal reflection and one-dimensional electron tunneling, but we eventually settled on a 1D photonic band gap as a more appropriate medium for tunneling. A dielectric mirror consists of alternating quarter-wave layers of high and low-index glasses, leading to constructive interference for reflection and destructive interference for transmission. Such a structure can be thought of as an analog of the Kronig–Penney model for a band gap in condensed-matter physics, and in fact there has been much work, both theoretical and experimental, on photonic band gaps (Yablonovitch [1993], John [1991]). The effective wave vector, or "quasimomentum", of light inside the band gap is imaginary, and we confirmed by direct numerical calculation that this qualitative similarity was sufficient to create the same saturating effect and superluminal transmission as tunneling through a rectangular barrier. It is important to note that there is no direct analog to the tunnel regime ($E < V_0$) for light; as shown in § 3, the analogy between the Schrödinger and Helmholtz equations leads to an effective index $n(x,y,z) = \{2m[E - V(x,y,z)]\}^{1/2}c/\hbar\omega$, which would be imaginary in any regions where $E < V$. Each microscopic (quarter-wave) region of the dielectric mirror is a region of allowed propagation, and it is only the Bragg reflection arising from the periodic spacing which makes the mirror as a whole a "forbidden region". The wave function can be written according to Bloch's theorem as a periodic Bloch function $u_k(\boldsymbol{r})$ times a plane wave $\exp\{\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}\}$; inside the band gap, k becomes imaginary, leading to an exponentially decaying field envelope, but $u_k(\boldsymbol{r})$ is still a sinusoidally oscillating function.

For our barrier, we chose an 11-layer mirror, with alternating indices of refraction of 1.41 and 2.22. At the design wavelength of 702 nm, this mirror had a transmission that dropped to about 1%; the band gap extended from 600 nm to 800 nm, over most of which range the group delay was smaller than $d/c = 3.6$ fs. The stationary phase approximation predicted that the group delay near midgap would saturate at approximately 1.7 fs. This structure had several other advantageous features. Unlike the microwave experiments, it involved negligible dissipation, and no dispersion outside the tunnel barrier. Furthermore, both the transmission probability and phase are very flat functions of frequency near midgap, so there is essentially no wave-packet distortion. Finally, the symmetry of the problem makes the semiclassical time vanish identically at midgap, which emphasizes that even that time cannot solve all causality problems, and allowed us to distinguish it quite easily from the group delay time.

Of course, direct electronic measurement of femtosecond-scale delays is not possible. We therefore used a nonlinear optical effect discovered by Hong, Ou and Mandel [1987], which can be thought of roughly as a time-reversed variant of the nonlinear autocorrelation technique for femtosecond laser pulses (which has also been applied to the tunneling problem by the Vienna group; see below). This effect relies on spontaneous parametric down-conversion, a process in which a crystal with a $\chi^{(2)}$ nonlinearity absorbs a pump photon at ω_0 and emits in its place a pair of photons (conventionally termed "signal" and "idler" despite the fact that in these experiments they are indistinguishable) at frequencies spread symmetrically about $\omega_0/2$, energy conservation being assured by the anticorrelation of the two photons' frequencies. The photons are emitted simultaneously to within their coherence lengths, and as the latter are only constrained by the phase-matching bandwidth and subsequent filters, one finds correlation times as short as 15 fs.

If the two photon wave packets meet simultaneously at opposite sides of a 50/50 beam splitter, a quantum interference effect related to Bose statistics causes them to exit the beam splitter along the same (randomly chosen) direction; detectors placed at the two exit ports of the beam splitter will never register photons simultaneously. On the other hand, if the two photons arrive at different times, each will make an independent choice at the beam splitter, leading to coincidence counts in half of the cases. Thus by changing the path length of one photon's trip until the coincidence rate is minimized, one can ensure that the photons are meeting simultaneously at the beam splitter (Hong, Ou and Mandel [1987], Steinberg, Kwiat and Chiao [1992b], Jeffers and Barnett [1993], Shapiro and Sun [1994]). If an obstruction such as a tunnel barrier is placed in one arm of the two-photon interferometer, the coincidence dip recorded as a function of external path length will shift, and this shift is a measure of the delay time for traversing the barrier. It is interesting to note that these experiments are typically performed with a continuous-wave argon laser as the pump, so the state of the light is in fact stationary in time. It is only the *correlations* between the photons which have the very fast (15 fs) time-dependence. Once a photon is detected, it is possible to say that its twin has "collapsed" into a 15-fs wave packet, but prior to that time, the system is better seen as a superposition of 15-fs wavepackets with centers at every possible position.

Other than the single-photon aspects, which were predicted theoretically not to modify the propagation times, this technique has some interesting advantages

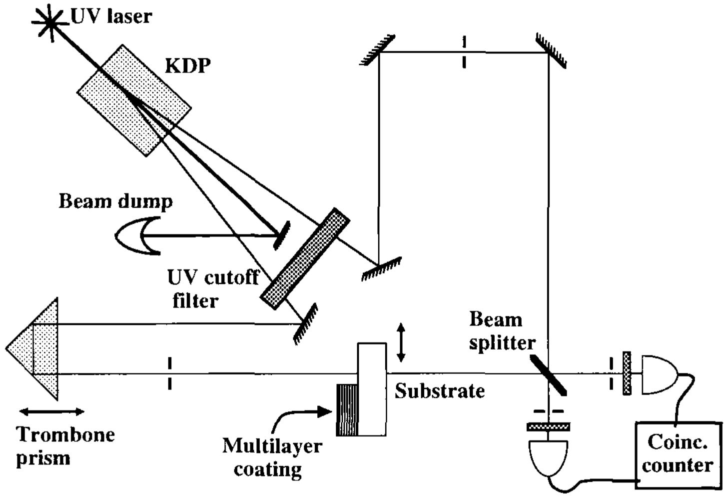

Fig. 6. Experimental setup for determining single-photon propagation times through a multilayer dielectric mirror.

relative to such approaches as classical white-light interferometry or nonlinear cross-correlation. Since the nonlinear effect is used only before the tunnel barrier, extremely low intensities may be used at the level of the sample; we typically counted on the order of 10^5 photons per second, by using tens of milliwatts of 351 nm light from an argon laser as a pump. As discussed by Steinberg, Kwiat and Chiao [1992a,b], Jeffers and Barnett [1993] and Shapiro and Sun [1994], first-order effects of group-velocity dispersion cancel out, allowing high resolution to be retained even in the presence of material dispersion. Finally, in contrast to standard interference techniques, this method relies *only* on detection of photon pairs, so the fringe visibility is not reduced by the low transmission through the tunnel barrier; interference occurs between two balanced Feynman processes, *each* of which involves only one tunneling event. Only the total count rate drops, leading to a $\sqrt{N}$ dependence for the uncertainty, which we countered by averaging a large number of 1-hour data runs.

By scanning across the coincidence dip while periodically inserting and removing the band gap coating (see fig. 6), we were able to measure the shift due to the propagation delay to better than 1 fs (Steinberg, Kwiat and Chiao [1993]). We also noted that as predicted, the shape of the coincidence dip (a direct measure of the overlap of the two wave packets) did not change significantly due to the presence of the barrier. In the first iteration of our experiment, we found the arrival time for propagation of a single photon through the 1.1 μm coating to

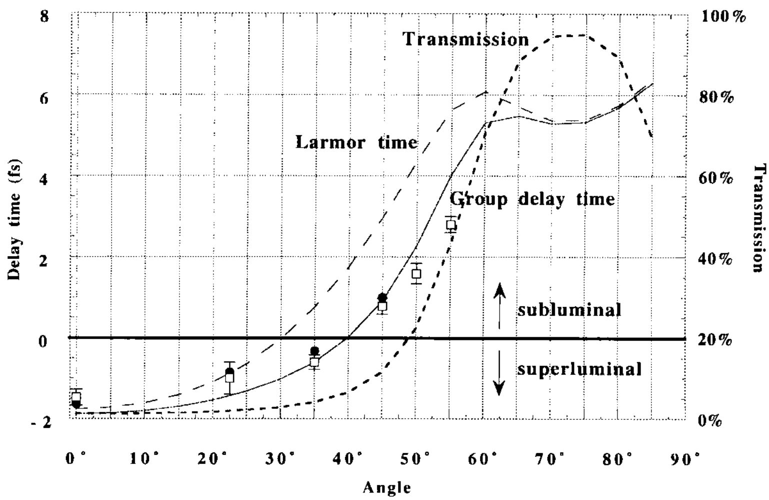

Fig. 7. Left axis: measured delay for mirrors 1 (squares) and 2 (circles) as a function of angle of incidence, to be compared with the group delay and with Büttiker's Larmor time. Right axis: transmission versus angle of incidence. All curves for p-polarization.

be *earlier* by 1.47 ± 0.21 fs than the arrival time for propagation through 1.1 μm of air. This 7-standard-deviation result confirmed the superluminality of single-photon tunneling. It would correspond to an effective tunneling velocity of $1.7c$. It differs from the stationary-phase prediction of 1.9 fs by about two standard deviations, and demonstrated immediately that the semiclassical time (which vanishes at midgap) was inadequate for describing wave packet propagation.

In a later extension of this experiment, we studied the frequency-dependence of the tunneling time (Steinberg and Chiao [1995]). Since it was not feasible to change the frequency of the photons in our interferometer, we changed the angle of incidence on the multilayer dielectric, thus altering the Bragg condition. In this way, we were able to scan from midgap nearly to the band-edge. We confirmed the qualitative behavior of the group delay, with absolute agreement generally better than 0.5 fs (see fig. 7). We were able to show that not only the semiclassical time but also Büttiker's Larmor time failed to describe the propagation effects[9].

[9] It is important to realize that these theories are not *intended* to describe propagation, but rather other aspects of tunneling. However, many researchers, made uncomfortable by the superluminal predictions of stationary phase, have expressed the expectation that these "interaction" times would in fact give the correct, *subluminal* time of arrival of a wave packet peak. Thus we did not disprove Büttiker's and Landauer's theories, but only demonstrated that their validity could not be extended to describe pulse propagation.

This refined data set showed a clearly visible change of sign of the pulse shift as the barrier was tuned from a regime of superluminal transmission to a subluminal one.

4.7. THE FLORENCE GROUP, PART II

In 1993, Ranfagni and co-workers, having become aware of the work of Ishii and Giakos, performed a new set of intriguing experiments (Ranfagni, Fabeni, Pazzi and Mugnai [1993]). They first repeated the latter's experiments on signal propagation in waveguides *above* cutoff, and found no evidence for any causality violation; while the phase velocity was indeed superluminal, the "signal" (their relatively abrupt step-modulated wave) travelled at the group velocity. They subsequently studied the claim of superluminal propagation in free-space. They measured a propagation speed of c for microwaves travelling between two horns which faced one another. When the receiver was translated perpendicularly to the propagation direction, however, they confirmed the surprising result that although the distance between the horns was increasing, the delay time displayed an initial decrease. In a mathematical analysis, they argued that this effect could be understood by analyzing the diffraction of the microwave out of the square aperture of the transmitter. The receiver was observing "leaky" evanescent waves in the shadow region of the near-field diffraction pattern. It is fascinating to note that the exponential decay of the field amplitude into this shadow region provides a qualitative analogy to tunneling. It begins to seem that exponential decay – whether due to absorption, tunneling, band gaps, or diffraction – leads in general to anomalous delay times. In the simplest cases, the imaginary wave vector is understood to lead to superluminal delays because no phase is accumulated along the propagation direction; in the newer examples where it is only an *envelope* which decays exponentially, the superluminality was not anticipated originally.

More recently, the Florentines have continued studying diffraction effects, this time using evanescent waves produced by a grating formed of metal strips (Mugnai, Ranfagni and Schulman [1997]). One of the evanescent modes was coupled through a paraffin prism onto a receiver (in analogy with the use of a second prism in frustrated total internal reflection). They have predicted that the group velocity will be superluminal in this case, as in the other examples of evanescent waves we have discussed. Experimentally, however, they were limited to measuring the phase shift at various frequencies, rather than performing a direct time measurement. They inferred the group velocity by numerically differentiating the resulting shift with respect to frequency (thus *assuming* the validity of the stationary-phase approximation), and the result they obtained suggested a time advance on the order of 50 ps over a distance of 3 cm, i.e., an effective velocity of about $2c$. These results, aside from being indirect, suffer from an amplification of the technical noise in the phase measurement. Ranfagni and co-workers are currently working on performing true time-dependent versions of this study.

4.8. THE COLOGNE GROUP, PART II

In 1994, the Cologne group extended their experiments to several new and interesting cases. Unfortunately, at the same time they extended their interpretational comments (which had been somewhat vague up to that point) to what could be interpreted as a nearly direct contradiction of Einstein causality, stating for example that "the superluminal propagation of frequency-limited signals by tunneling modes is possible". In order to sharpen up the debate over the meaning of signal propagation (somewhat clouded in much of the literature by the consideration of admittedly idealized situations involving infinitely high-frequency components and analytic wave forms), they encoded Mozart's 40th Symphony on a microwave signal which they claimed subsequently to have transmitted at $4.7c$.

Since many of these disputes frequently boil down to semantics, and since the workers involved have nonetheless found it impossible to find working definitions which removed all disagreement, it is perhaps best to quote the Cologne group directly (Heitmann and Nimtz [1994]): "The signals considered in the microwave experiments were unlimited in time and not Gaussian. Therefore Enders and Nimtz have never claimed that the front of a signal has travelled at superluminal speed. However, they have stated that the peak and the rising edge of a frequency band limited wave packet propagate faster than c through a barrier. This result corresponds to a superluminal group and signal velocity and it was recently used to transmit Mozart's Symphony No. 40 through a tunnel of 114 mm length at a speed of $4.7c$".

In fact, as will be seen below in our discussion of causality and superluminality (see § 8), this appearance of a wave form faster than c is in itself nothing surprising. This becomes particularly clear when one considers the timescales involved. The time advance being discussed is well under 1 ns in Nimtz's experiments. An acoustic wave form, on the other hand, has a useful bandwidth on the order of 20 kHz, which is to say that no significant deviation from a low-order Taylor expansion occurs in less than about 50 µs. To predict where the wave form would be 50 µs in advance requires little more than a good eye; to predict it 1 ns in advance hardly even requires a steady hand. As was already suggested by Chiao, Kwiat and Steinberg [1993] and Steinberg [1994, 1995c], and recently made more explicit by Kurizki and Japha [1993] and Diener [1996], the interference at work in tunneling has the effect of advancing the incident wave form due to the first derivative term of Taylor's theorem [10]. Hence even though the transmitted wave mimics the future behavior of the incident wave impressively well, it does so without any need for *information* about the later behavior of the incident field. The already existing information at any given time is more than sufficient to make an educated guess about what is to come a short time later, and a tunnel barrier does no more than act as an analog computer for this purpose. All the same, this ability (particularly when coupled with amplification, as will be discussed below) does provide an interesting way to advance the triggering of a fixed-discriminator-level detection system, and may not be without technological application. Of course, it becomes even more surprising when we are not merely arguing about the shape of a classical wave form, but the unique time of arrival (i.e., the "click") of an individual quantum particle. Since this latter quantity is tied inextricably to interpretational issues (such as the frequently invoked "instantaneous" collapse), no solution is likely to be forthcoming soon.

Leaving aside these interpretational issues for the moment, the recent series of experiments in Cologne extend the microwave work to new barriers, including an analog of the periodic-dielectric structure first studied at Berkeley. Although some of them rely again on phase measurements, and the signal-to-noise ratio remains dubious, they provide an elegant confirmation, and reach effective speeds of several times that of light. Furthermore, Nimtz and co-workers have been able to verify again the thickness-independence of the tunneling time in the opaque limit. Finally, since microwave experiments are plagued by effects of dissipation in the waveguides, they have performed interesting studies on tunneling in the presence of dissipation, which has also been analyzed in various other frameworks (Nimtz, Spieker and Brodowsky [1994], Mugnai, Ranfagni, Ruggeri and Agresti [1994], Raciti and Salesi [1994], Steinberg [1995b], Brodowsky, Heitmann and Nimtz [1996]).

[10] If destructive interference is set up between part of the wave travelling unimpeded and part which has suffered a small delay Δt due to multiple reflections, one has $\Psi_{\text{out}}(t) = \Psi_{\text{in}}(t) - \xi\Psi_{\text{in}}(t-\Delta t) \approx (1-\xi)\Psi_{\text{in}}(t) + \xi\Delta t\, \mathrm{d}\Psi_{\text{in}}(t)/\mathrm{d}t \approx (1-\xi)\Psi_{\text{in}}(t+\xi\Delta t/(1-\xi))$, which is already a linear extrapolation into the future. In cases where the dispersion is sufficiently flat, as in a bandgap medium, the extrapolation is in fact surprisingly better than this first-order approximation. As was suggested in Steinberg [1995c] and recently discussed more rigorously by Lee and Lee [1995] and Lee [1996], this implies that even a simple Fabry–Perot interferometer exhibits superluminality when excited off resonance.

4.9. THE VIENNA GROUP

The Berkeley work, in which multilayer dielectric mirrors functioned as photonic band gap media and hence as effective tunnel barriers, was extended by the ultrafast laser group at the Vienna Technical University in 1994. By using 12-fs laser pulses and standard nonlinear-optical autocorrelation techniques, they benefitted from a better signal-to-noise ratio than the single-photon counting experiments, and were therefore able to study barriers of lower transmission. Of course, in so doing, they were only able to study classical electromagnetic pulses, disregarding the single-particle features, but as we have discussed, the single-photon arrival times had been seen to be quite well described by Maxwell's equations. Since such group-delay measurements are incapable of addressing deeper issues of particle–wave duality (for these, "clocks" such as the Larmor clock to be discussed further below are essential), the sacrifice is not a great one.

Spielmann, Szipöcs, Stingl and Krausz [1994] used 12-fs FWHM sech-squared optical pulses with energies of about 1 nJ at a repetition rate near 100 MHz to measure transmission times through quarter-wave stacks of 6, 10, 14, 18, and 22 layers, with transmissions ranging from 30% to $2 \cdot 10^{-4}$ (compare the 11-layer Berkeley structure with its 1% transmission, near the noise limit for that experiment). A freely-propagating pulse was compared with one which had to traverse the coating being studied, and the two pulses were subsequently superposed in a non-collinear geometry in a BBO crystal to generate second-harmonic light and thus a background-free cross-correlation signal.

Since the required time resolution was of the order of 1 fs, while the bandwidth-limited pulses were 10 to 15 times longer, a multishot averaging technique was used. This requires extremely high stability of the pulse parameters, which the Vienna group achieved thanks to a mirror-dispersion-controlled Ti:sapphire laser (Stingl, Spielmann, Krausz and Szipöcs [1994]). This laser generated bandwidth-limited pulses at 800 nm, with close to 1% stability in the frequency-doubled output.

They split each pulse in two parts, which were superposed in the nonlinear crystal after one part traversed the dielectric coating while the other propagated in air. The cross-correlation signal varied as a function of the degree of overlap of the two pulses in the crystal. By adjusting the path-length difference to put themselves on the edge of the output signal, and then switching the coating between the two arms of the correlator, the researchers were able to measure small shifts in the pulse position caused by the coating (Spielmann, Szipöcs, Stingl and Krausz [1994]). Great care was taken to eliminate systematic errors due to the change in the shape of the cross-correlation signal occasioned by the insertion and displacement of the sample, due to drifts and fluctuations of pulse parameters, etc. The experimenters thus obtained results with statistical uncertainties of ± 0.3 fs, and by studying progressively thicker samples, they were able to confirm the prediction that the time delay should saturate at a finite value even as the thickness of the sample continued to grow. For the thickest sample studied, they found an advance of about 6 fs over free propagation in air. However, their results showed a systematic deviation from the stationary-phase prediction of about one and a half femtoseconds; this discrepancy is not yet understood. They did observe that the 28-THz bandwidth pulses from their laser were somewhat distorted, at least by the 22-layer barrier, based on interferometric autocorrelation traces. The pulse width decreased from 12 fs to 6.5 fs, consistent with the effectively increased bandwidth due to the lower transmission at the center of the pulse spectrum than in the wings. Since the wings also have a longer group delay than the center frequency, it is possible that the observation of slower-than-predicted traversal is in part due to the preferential transmission of these slower components, but a full explanation has not yet been given.

4.10. DEUTSCH AND GOLUB'S LARMOR-CLOCK EXPERIMENT

Deutsch and Golub [1996] performed an experiment to measure the Larmor tunneling time for photons. Their experiment utilizes an analogy between the spin of an electron and the spin of a photon, whose polarization state can be described by a point on the Poincaré sphere given by the Stokes parameters $\boldsymbol{S}$. The equation of motion for the Stokes parameters for a beam of light propagating along the x-axis through a medium with an anisotropic refractive index is given by

$$\mathrm{d}\boldsymbol{S}/\mathrm{d}x = \boldsymbol{\Omega}\times\boldsymbol{S}, \tag{4.3}$$

where $\boldsymbol{\Omega}$ is the precession rate of the tip of the $\boldsymbol{S}$ vector on the Poincaré sphere arising from the anisotropic index of refraction. This equation is formally identical to the one describing the precession of the tip of the electron spin vector $\boldsymbol{\sigma}$ on the Bloch sphere arising from an applied magnetic field

$$\mathrm{d}\boldsymbol{\sigma}/\mathrm{d}t = \boldsymbol{\Omega}_{\mathrm{L}}\times\boldsymbol{\sigma}, \tag{4.4}$$

when the optical precession rate $\boldsymbol{\Omega}$ is identified (apart from a proportionality constant) with the rate of Larmor precession $\boldsymbol{\Omega}_{\mathrm{L}}$.

This analogy between electron and photon spin precession led Deutsch and Golub to suggest an optical implemention of the Larmor-clock measurement of the tunneling time of Baz' and Rybachenko (latter corrected and generalized by Büttiker). The basic idea is to replace electrons with photons, and to replace a uniform magnetic field confined to the electron tunnel barrier region with a uniform birefringent medium confined to the corresponding optical tunnel barrier. Thus, instead of utilizing the precession of the electron spin as an internal clock to measure the Larmor tunneling time, they utilized the precession of the $\boldsymbol{S}$ vector of the photon as an internal clock. In their experiment, they used frustrated total internal reflection between two glass prisms as the tunnel barrier. The gap between the prisms, which served as the tunnel barrier, was filled with a birefringent fluid (a liquid crystal).

There are a number of advantages in performing an experiment using photons to measure the Larmor tunneling time. In contrast to the case of electrons, it is easy to confine the region for photon spin precession to the region of the barrier, by simply restricting the birefringent fluid to the region of the gap, whereas it is hard to confine the magnetic field to the region of the tunneling barrier for electrons. Also, since the photon is neutral, complications inherent in electron tunneling-time measurements associated with image charges induced in the faces of the tunnel barrier could be avoided. Moreover, the interaction between the photons is negligibly weak, in contrast to the strong Coulomb repulsion between the electrons inside the barrier. Exploiting these advantages, Deutsch and Golub successfully completed their experiment to measure the Larmor tunneling time, with the result that the theoretical predictions of Büttiker for the Larmor time were qualitatively confirmed.

However, in a critical examination of their own experiment, Deutsch and Golub pointed out a weakness: the Larmor tunneling time is based ultimately on an arbitrary definition that is, in their words, "not a physical scale that emerges naturally, or that is needed to calculate the results of measurements". They pointed out another possible weakness: the process as measured by the Larmor clock is a stationary one involving only a single energy or frequency of the photon. It has been argued that the tunneling time cannot have any meaning for *stationary* processes, which have no beginning or ending (Falck and Hauge [1988], Gasparian and Pollak [1993], Gasparian, Ortuño, Ruiz, Cuevas and Pollak [1995], Krenzlin, Budezies and Kehr [1996]). However, we shall see that stationary processes *can* in fact give *indirect* information on tunneling times in 2D situations, as has been demonstrated by the continuous-wave experiments of Balcou and Dutriaux (see the next section).

In her PhD thesis, Deutsch gave a theoretical treatment of the *nonstationary* problem of an *interacting* photon system inside the tunneling barrier, interacting via a third-order nonlinear optical susceptibility confined to the region of the barrier. The basic idea is that when one photon is inside the barrier region, it causes a refractive-index change through the nonlinear susceptibility, which tends to exclude (for the repulsive sign of the nonlinearity) the presence of a second incident photon which is about to enter the barrier. The tunneling time was defined as the duration over which the second photon tends to be excluded by the first photon. Thus one could determine the tunneling time through Glauber's two-photon correlation function, as applied to a nonlinear beam splitter used as a model for the tunneling barrier. The result of the calculation was a certain correction term in the two-photon correlation function which arose from the nonlinearity. She made an identification of the resulting tunneling time with the dwell time. However, as many workers have pointed out (Hauge and Støvneng [1989]), the dwell time cannot distinguish between reflected and transmitted particles, and hence cannot be regarded as a genuine tunneling time; we will see in § 5 how one might hope to get around such objections.

4.11. BALCOU AND DUTRIAUX'S FTIR EXPERIMENT

Tunneling times have been measured recently in frustrated total internal reflection (FTIR) by Balcou and Dutriaux [1997]. The idea of this beautifully simple experiment is to utilize both the lateral displacement and the angular deflection of the transmitted light beam (which is composed of the tunneling photons), as a simultaneous measurement of two different kinds of tunneling times, which turned out to be the group delay and the semiclassical time. These two tunneling times correspond to the real and imaginary parts of a complex time related closely to that of the Larmor times of eq. (2.9). In § 5, we shall see that it is possible to delineate clearly the physical significances for these two different times.

Let us define the x-axis as the direction normal to the interface between the prisms and y-axis as the direction parallel to the interface in the plane of incidence (see fig. 8). This 2D FTIR tunneling geometry has been analyzed previously by Steinberg and Chiao [1994a] and by Lee and Lee [1997]. During the tunneling process which occurs in the x-direction, the wave packet continues to propagate in the y-direction, since its y-component of momentum is conserved. Balcou and Dutriaux argue heuristically that one expects the propagation velocity along the y-axis to be uniform during tunneling, and that, therefore, this would result in a lateral displacement D along the y-direction which would be proportional to some unknown temporal delay due to tunneling.

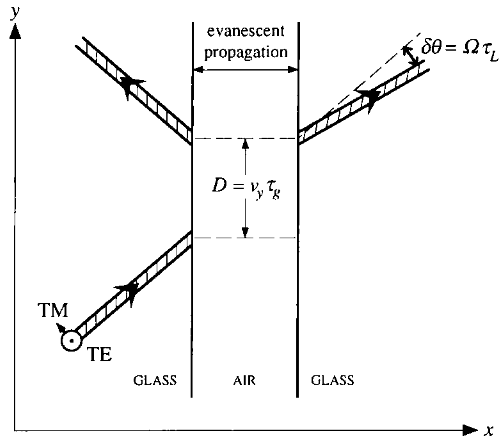

Fig. 8. Schematic of Balcou and Dutriaux's frustrated total internal reflection experiment to measure two tunneling times. These two times are inferred from the Goos–Hänchen shift D, and from the angular deflection of the transmitted beam $\delta\theta$, respectively.

Calculations (Ghatak, Shenoy, Goyal and Thyagarajan [1986], Ghatak and Banerjee [1989]) show that after the wave packet has finished tunneling through the interface in the opaque limit, it is the group delay τ_g which causes the lateral displacement of the transmitted wave packet along the y-direction by an amount $D = v_y \tau_g$, where $v_y = c/n \sin\theta$ is the y-component of the velocity of the wave packet (θ being the angle of incidence). This lateral shift of the transmitted light beam turns out to be identical to the well-studied Goos–Hänchen shift. Therefore, Balcou and Dutriaux infer that a measurement of the displacement D will lead to the tunneling time

$$\tau_g = D[c/n \sin\theta]^{-1}. \tag{4.5}$$

In addition to this lateral displacement, there is also an angular deflection of the transmitted beam, which arises from its finite beam size. Due to diffraction, the finite width of the incident beam of light leads to some finite spread in the angles of its wave vectors. Larger angles are transmitted less than smaller angles, since they are farther away from the critical angle. This causes a preferential transmission of the smaller angle components of the incident beam, which leads to a deflection of the transmitted beam slightly towards the normal. This is analogous to the effect associated with Büttiker's Larmor time in which there is a preferential transmission of electron spins aligned antiparallel to the magnetic field, which leads to a slight spin polarization of the transmitted beam.

Calculations similar to those above show that this preferential transmission leads to an effective angular frequency of rotation of the beam at the rate

$$\Omega = nc/2w_{\mathrm{R}} \csc(2\theta), \tag{4.6}$$

where w_{R} is the beam Rayleigh length. Balcou and Dutriaux therefore infer that a measurement of the angular deflection $\delta\theta$ will yield the tunneling time

$$\tau_{\mathrm{L}} = \delta\theta/\Omega, \tag{4.7}$$

where τ_{L}, the so-called "loss time", approaches the semiclassical time of Büttiker and Landauer for opaque barriers. The two tunneling times τ_{g} and τ_{L} turn out to be identical to the real and imaginary parts of the complex tunneling time introduced by Pollak and Miller [1984],

$$\tau_c = \tau_{\mathrm{g}} + \mathrm{i}\tau_{\mathrm{L}} = -\mathrm{i}\frac{\partial \ln t}{\partial\omega}, \tag{4.8}$$

where t is the complex transmission coefficient of the tunnel barrier.

Balcou and Dutriaux obtained experimental data which agree well with the above theory for the two tunneling times. In particular, they have demonstrated not only that the group delay saturates with increasing barrier thickness (the Hartman effect), but also that the semiclassical time increases linearly with this thickness. However, they interpreted the semiclassical time as the one "most relevant to describe the physics of tunneling", in contrast to the group delay. They do so for two reasons. First, the semiclassical time "yields only subluminal velocities so that the causality principle is explicitly obeyed", in contrast to the group delay, which yields superluminal velocities. Second, the group delay is dependent on the boundary conditions, and differs considerably for TM and TE polarized light, whereas the semiclassical time is independent of these boundary conditions. They argue that since a tunneling time should be independent of boundary conditions (it should depend only on what happens in the *interior* of the barrier), this singles out the semiclassical time as the true tunneling time.

In answer to their first point, in point of fact the semiclassical time under certain circumstances can also be superluminal, a point which they failed to recognize. In the case of the 1D photonic band gap discussed earlier, the semiclassical time is zero at midgap (Martin and Landauer [1992], Steinberg, Kwiat and Chiao [1993]), which is a behavior even *more* superluminal than that predicted by the group delay for this kind of barrier.

In answer to their second point, boundary conditions are in fact very important for tunneling. Again, in the example of evanescent waves in the 1D photonic band gap, it is the Bragg reflections from the periodic dielectric boundaries which give rise to the band gap, and hence tunneling. These reflections would of course vanish if there were no boundary conditions necessary for the partial reflections at the interfaces between the successive dielectrics, and tunneling would disappear. More generally, tunneling is a wave-interference phenomenon. Since boundary conditions are important for determining this interference, it is unreasonable to demand that the tunneling time be independent of boundary conditions. Hence, as should become clear in the following section, we disagree with their conclusion that it is the semiclassical time, not the group delay, that is the one "related solely to tunneling". Rather, we believe that their results constitute experimental evidence for the simultaneous existence of these two tunneling times in the same barrier.

§ 5. New Theoretical Progress

One commonly cited reason for the difficulty of defining a tunneling time unambiguously is the fact that time in quantum mechanics does not have the status of a Hermitian operator, and can thus not be measured directly. This is not an airtight objection, since most physical measurements are in fact indirect: we say we have measured the position of a particle when what we may in fact have observed is which element of a CCD array absorbed photons scattered by the particle and then focused. Even in classical mechanics, one never measures "the time of a particle", or even "the time of an event", but a quantity such as the angle through which a stopwatch hand rotates if it is started by the particle's entry into a region and stopped by its exit from that region. When many different operational definitions of this sort yield the same result, we feel justified in calling the quantity we have found "*the* time"; if, as in the tunneling case, different measurements yield different results, we must be more cautious.

In quantum mechanics, it is straightforward to define an operator Θ_B which is 1 if the particle is in the barrier region and 0 otherwise. Such a projection operator is Hermitian, and may correspond to a physical observable. Its expectation value simply measures the integrated probability density over the region of interest– it is this expectation value divided by the incident flux which is referred to as the dwell time. Thus the central problem is not the absence of an appropriate Hermitian operator, but rather the absence of well-defined histories (or trajectories) in standard quantum theory. For example, the dwell time measures a property of a wave function with both transmitted and reflected portions, and does not display a unique decomposition into portions corresponding to these individual scattering channels. Some workers calculate the expectation value not for the initial state but rather for the final state (van Tiggelen, Tip and Lagendijk [1993]). This answers the question no better than does the usual dwell time; instead of discarding information about late times, it discards information about early times. Approaches relying on projector algebra in general have been analyzed by Muga, Brouard and Sala [1992b] and Leavens [1995]. Other related approaches follow phase space trajectories (Muga, Brouard and Sala [1992a]), Bohm trajectories (Dewdney and Hiley [1982], Leavens [1990, 1993], Leavens and Aers [1991, 1993], Leavens, Iannaccone and McKinnon [1995], Leavens and McKinnon [1995]), or Feynman paths (Sokolovski and Baskin [1987], Sokolovski and Connor [1990, 1993, 1994], Hänggi [1993], Fertig [1990, 1993]). No consensus has been reached as to the validity and the relationship of these various approaches. Ideally, transmission and reflection times τ_T and τ_R would, when weighted by the transmission and reflection probabilities $|t|^2$ and $|r|^2$, yield the dwell time τ_d:

$$|t|^2\tau_T + |r|^2\tau_R = \tau_d; \tag{5.1}$$

this relation has served as one of the main criteria in a broad review of tunneling times (Hauge and Støvneng [1989]), but has also been criticized (see, for example, Landauer and Martin [1994]).

However, a formalism due to Aharonov, Albert and Vaidman [1988] and Aharonov and Vaidman [1990] shows how to analyze "conditional measurements" in quantum mechanics; that is, how to predict outcomes of measurements not for entire ensembles, but for *sub*ensembles determined both by state preparation and by a subsequent postselection. In the case which concerns us, the state is prepared with a particle incident from the left, and selected to have a particle emerging on the right at late times. Due to the time-reversibility of the wave equation, results of intervening measurements depend both on the initial and the final state. This formalism relies only on standard quantum theory, and yields a result that is completely general for any measurement arising from a von Neumann-style measurement interaction, in the limit where the interaction strength is kept low enough to avoid irreversibly disturbing the quantum evolution. This low strength implies great measurement uncertainty on any individual shot, but an average may be calculated for a large number of data runs. We have recently shown (Steinberg [1995a,b]) how to apply this formalism to tunneling, and the time we find is identical to the complex time of Sokolovski, Baskin, and Connor, τ_c. But thanks to the "weak measurement" formalism, it becomes clear what the physical significance of the real and imaginary parts is: the real part (the in-plane Larmor time) quantifies how strongly the tunneling particle will affect a clock with which it interacts; this is the portion which corresponds to a classical measurement outcome. The imaginary part, on the other hand, describes the amount of back-action the measuring apparatus will exert on the particle (the sensitivity of the tunneling probability to small perturbations, in other words, as in Büttiker's out-of-plane Larmor rotation). While the former effect remains constant as the measurement is made weaker and weaker, the back-action may be made arbitrarily small by resorting to extremely "gentle" (and consequently uncertain) measurements. Among other attractive properties, these conditional times automatically satisfy eq. (5.1).

The generality of the times obtained in this way suggests that it may be possible to apply them to a broad variety of problems, at least approximately, even in cases where exact solution would be intractable. It has already been shown that not only are the Larmor times a clear subset of these "conditional times", but that the counter-intuitive effects of absorption on light propagating through layered media can be understood qualitatively by application of these complex times (Steinberg [1995b]). The equivalence of τ_{BL} and $-\mathrm{Im}\,\tau_c$ makes sense given that the oscillating-barrier approach in fact studies the sensitivity to perturbations in the barrier height. The direct connection to measurement outcomes lifts the ambiguity present in other "projector approaches" and the Feynman-path formalism. Finally, it is possible using these methods to calculate conditional probability distributions for transmitted or reflected particle positions as a function of time, and to directly investigate questions about whether tunneling particles spend significant lengths of time in the center of the barrier, whether only the leading edge of the wave is transmitted, etc. Since these probability distributions may have large values on both sides of the barrier simultaneously, and independent "weak measurements" can be shown to add linearly (unlike "strong" measurements of non-commuting observables), it is interesting to speculate about whether a statistical demonstration that during tunneling, a particle is "in two places at once" might be possible. Work continues on all of these issues. Extensions are also underway to analyze whether one can go a step beyond these expectation-value-like tunneling times and calculate higher moments, or entire distributions (Iannaccone [1996]).

§ 6. Tunneling in de Broglie Optics

Tunneling was, of course, discussed *per se* for electrons before the analogy to optical effects was drawn. However, it is an effect that is quite general to wave propagation. Future promising directions for studying tunneling rely on a variety of particles and barriers with their own particular advantages and difficulties. Recently, workers at Kyushu University and the Research Reactor Institute in Osaka have used a neutron spin echo instrument to measure Larmor precession (and thus Larmor times) for neutrons traversing a magnetic layer (Hino, Achiwa, Tasaki, Ebisawa, Akiyoshi and Kawai [1996]). Preliminary results appear to agree well with theory, even near the critical angle for total reflection of the neutrons, and there is every reason to expect more interesting data to come from studies of neutron tunneling.

Ballistic transport and even refraction of electrons in heterostructures has been described theoretically (Gaylord, Henderson and Glytsis [1993]) and observed experimentally (Spector, Stormer, Baldwin, Pfeiffer and West [1990]). It is clearly feasible to extend these geometries and observe frustrated total internal reflection of electrons. As discussed by Steinberg and Chiao [1994a], there is a number of interesting similarities and differences between tunneling of massive and massless particles and between one- and two-dimensional tunneling. Future studies with ballistic electrons ought to be able to shed new light on aspects of the tunneling problem (Lee and Lee [1995]). They will also be closer to areas which are likely to be of technological impact (Spector, Stormer, Baldwin, Pfeiffer and West [1990], for example, have demonstrated a new kind of electronic switch relying on electron refraction).

Atoms also display wave properties. For a number of years now, atom interferometers have been in operation, and recently both Bose–Einstein condensation and a coherent pulsed output coupler for such matter waves have been observed (Anderson, Ensher, Matthews, Wieman and Cornell [1995], Mewes, Andrews, Kurn, Durfee, Townsend and Ketterle [1997], Andrews, Townsend, Miesner, Durfee, Kurn and Ketterle [1997]). The tunneling of such composite particles is in a sense even more striking than that of photons, neutrons, or electrons. The wealth of internal degrees of freedom of an atom also makes it an attractive candidate for studying a variety of "interaction times". With the latest laser-cooling and -trapping techniques, atoms may now be produced with de Broglie wavelengths significantly larger than an optical wavelength, meaning that tunnel barriers can be constructed from tightly focussed light beams, making use of the repulsive dipole force (Steinberg, Thompson, Bagnoud, Helmerson and Phillips [1996]). Auxiliary probe beams interacting with the atoms while in or near the tunnel region could be used to make the atoms fluoresce (Japha and Kurizki [1996a]), or to optically pump them, or (in order to avoid any dissipation) to induce Raman transitions. By looking at atoms transmitted through such beams, at Toronto we plan to study a number of interaction times, as well as their position- and spatial-dependence, as discussed by Steinberg [1995a,b]. Multiple simultaneous probe beams would allow one to investigate further issues of locality and the "reality of the wave function". We are also studying the conjecture that position-dependent magnetic fields, which can rapidly tune atoms through either Raman or RF resonances (which can be extremely narrow on the scale of feasible Zeeman shifts even over length scales much smaller than an optical wavelength (Thomas [1994])), can be used to create extremely thin interaction regions which will lead to quantum reflection and tunneling once the de Broglie wavelength is longer than the interaction length (cf. Kurizki [1997]). Such mechanisms would allow even more sensitive studies, as well as extensions to more complicated geometries, such as thin Fabry–Perot cavities for atoms.

Tunneling of atoms has already been observed in a very different context. Investigating the behavior of ultracold atoms in a standing wave, Raizen's group at the University of Texas has observed a number of fascinating effects related to the band structure of the atoms' center-of-mass motion in a periodical optical potential, including the analog of Landau–Zener tunneling when the optical potential is accelerated fast enough that the atoms begin to tunnel to a higher band (Niu, Zhao, Georgakis and Raizen [1996]).

§ 7. Superluminality and Inverted Atoms

The fact that superluminal wave packet propagation through tunneling barriers has been observed experimentally leads naturally to the following question: Are there any other situations in physics where such superluminal behavior can arise? Of course it would be nonsensical to ask: Can light go faster than light? But it does make sense to ask the question: Can light *in a medium* go faster than light *in the vacuum*? Surprisingly, the answer to this question is "yes" in at least one instance other than in tunneling, namely, when off-resonance pulses propagate through a medium with inverted atomic populations; that is, when wave packets are tuned to a *transparent* spectral region *outside* of the gain line (Chiao [1996]).

There are two situations in which closely related superluminal propagation effects appear in media with atomic population inversion. In the first situation, a *steady-state* one, an index of refraction model of the medium leads to an accurate description of the behavior of the system. When a two-level system is pumped *steadily* so that it becomes inverted, the real part of the linear susceptibility of the inverted two-level medium suffers a sign change relative to that of an uninverted medium, leading to superluminal group velocities in transparent spectral windows far away from resonance (Chiao [1993]). In the second situation, a *transient* one, the *sudden* inversion of the two-level system by a pulsed pump gives insight into the collective behavior of the system. Undamped atomic polarization waves are coupled strongly to electromagnetic waves, and this coupling leads to tachyon-like collective excitations, i.e., normal modes of the coupled atom-radiation system which exhibit a tachyon-like dispersion relation near resonance (Chiao, Kozhekin and Kurizki [1996]). It should be noted at the outset that these situations will lead to superluminal propagation phenomena which are much more dramatic than those which occur in tunneling, since no appreciable attenuation or reflection of the wave packets will occur in these dilute, *transparent* media, and consequently the distances over which superluminal propagation occurs can be much larger than those that occur in tunneling barriers.

As an example of the first, *steady-state* situation, we shall focus on the special case of superluminal propagation of finite-bandwidth pulses through a population-inverted medium, whose carrier frequencies are much lower than resonance. Although superluminal propagation also occurs near the resonance line[11], it is much simpler to understand the very-low-frequency case first. The refractive index of a two-level medium can be obtained from the usual Lorentz model, which yields (Jackson [1975], Kittel [1986])

$$n(\omega) = \left(1 + \frac{\omega_p^2}{\omega_0^2 - \omega^2 - i\gamma\omega}\right)^{1/2}, \tag{7.1}$$

where γ is a (small) phenomenological linewidth, ω_0 is the resonance frequency of the medium, and ω_p is "the effective plasma frequency", a measure of the strength of the coupling between the atoms and the radiation field, which is given by

$$\omega_p = (-4\pi w f N e^2/m)^{1/2}. \tag{7.2}$$

The Lorentz model has been generalized to include the possibility of population inversion, based on the density-matrix equations of motion for the two-level atom (Boyd [1992]), by introducing into eq. (7.2) the fractional atomic population inversion w, which is given by

$$w = \frac{N_u - N_l}{N_u + N_l}, \tag{7.3}$$

N_u being the number density of atoms in the upper level, N_l being the number density of atoms in the lower level, and $N = N_u + N_l$ being the total number

[11] An experiment is presently being performed at Berkeley using the stimulated Raman effect in rubidium vapor to demonstrate these resonantly enhanced superluminal group velocities (Chiao [1994], Chiao, Bolda, Bowie, Boyce, Garrison and Mitchell [1995]).

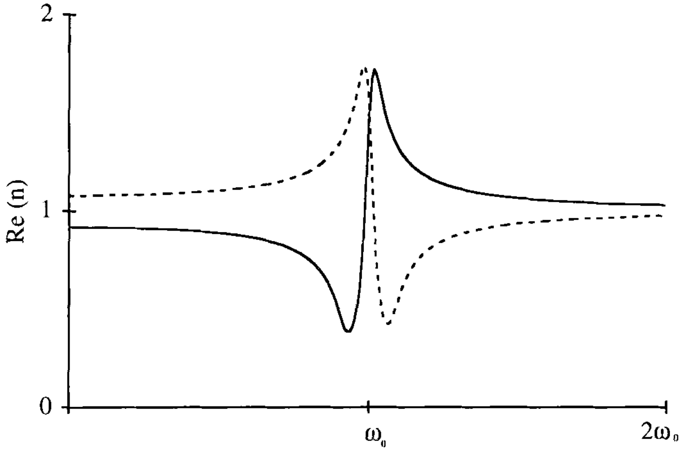

Fig. 9. Real part of the refractive index versus frequency for a completely inverted two-level atomic medium (solid line for $w = +1$), compared with that for the same medium with completely uninverted populations (dashed line for $w = -1$).

density of atoms in the two-level system. As usual, e is the electron charge, and m is the electron mass. The single-atom oscillator strength of the transition between these two levels is given by

$$f = 2m\omega_0 |\langle u|x|l\rangle|^2/\hbar = 2m(E_u - E_l)\,|\langle u|x|l\rangle|^2/\hbar^2 \tag{7.4}$$

where E_u and E_l are the energies of the upper and lower states of the atom, respectively, and $\langle u|x|l\rangle$ is the transition matrix element between these two states. In the special case when all the atoms are in the lower level ($w = -1$), the effective plasma frequency is real, but when there is complete population inversion and all the atoms are in the upper level ($w = +1$), the effective plasma frequency becomes imaginary. When one completely inverts the system, the inversion process can be thought of as an interchange of the two energy levels of the atom E_u and E_l, thus leading effectively to a sign change in the oscillator strength given by eq. (7.4). Thus for each atom, $f \rightarrow -f$ upon a complete inversion of the system.

Now let us consider the typical situation in which the inequalities $\gamma \ll \omega_p \ll \omega_0$ are obeyed. A plot of the real part of eq. (7.1) is shown in fig. 9. The extreme case of $w = -1$, with all the atoms in the lower level, where there is maximum absorption, is represented by the dashed line, and the opposite extreme case of $w = +1$, with all the atoms in the upper level, where there is maximum gain, is represented by the solid line. Note that the nature of the dispersion has been reversed for these two cases – regions of normal dispersion are interchanged with regions of anomalous dispersion upon the inversion of atomic populations, and vice versa. Also note that the sign in front of the second term under the square root in eq. (7.1) reverses upon population inversion – it is positive for the uninverted medium, but it becomes negative for the inverted medium. The physical meaning of this second term is that it represents the complex, frequency-dependent susceptibility of the medium (apart from a constant of 4π). This complex susceptibility reverses sign upon an inversion of population. Hence the imaginary part of the susceptibility reverses sign, which indicates the passage of the system from absorption into amplification. The real part of the susceptibility also reverses sign (Chiao and Boyce [1994]), which indicates the passage of the system from subluminality into superluminality in transparent spectral regions far away from resonance. In particular, as a result of this sign change, the index of refraction near zero frequency passes from a value greater than unity, over to a value less than unity given by

$$n(0) = (1 - |\omega_p|^2/\omega_0^2)^{1/2} < 1 . \qquad (7.5)$$

This result is valid whenever a strong, low-frequency resonance dominates the zero-frequency sum rule, e.g., when there exists an inverted population in the 24 GHz ammonia resonance used in the first maser (Chiao [1996]).

From eq. (7.1) it also follows that the slope $d[\mathrm{Re}\, n(\omega)]/d\omega$ approaches zero as $\omega \to 0$. Since the resulting group velocity dispersion vanishes near zero frequency, the medium is essentially dispersionless near DC (see fig. 9), a fact which is true for both the inverted and the uninverted media.

Now consider the propagation of a classical, finite-bandwidth pulse, for example, a Gaussian wave packet, whose carrier frequency and spectrum lie far below the resonance frequency of the two-level atom. Let this wave packet be incident upon a population-inverted medium. The amplitude of this wave packet will be chosen sufficiently small so that only the *linear* response of the medium to this weak perturbation need be considered.

The fact that the index $n(0) < 1$ is less than unity means that the phase velocity

$$v_p(0) = c/n(0) > c \qquad (7.6)$$

is greater than the vacuum speed of light c. It is well known that the phase velocity can exceed c without any violation of special relativity. (The phase velocity, which is the velocity of the zero-crossings of the carrier wave, characterizes the motion of a pattern which carries no information with it.)

More surprisingly, here as zero frequency is approached, the group velocity

$$
\begin{aligned}
\upsilon_{\mathrm{g}}(0) &= \left(\frac{\mathrm{d}\,\mathrm{Re}\,k(\omega)}{\mathrm{d}\omega}\right)^{-1}_{\omega\to 0} \\
&= c\left[\mathrm{Re}\,n(\omega) + \omega\frac{\mathrm{d}\,\mathrm{Re}\,n}{\mathrm{d}\omega}\right]^{-1}_{\omega\to 0} \\
&= \frac{c}{n(0)} = \upsilon_{\mathrm{p}}(0) > c
\end{aligned}
\tag{7.7}
$$

is equal to the phase velocity, and is therefore also superluminal: *The group velocity also exceeds the vacuum speed of light.* Furthermore, there is negligible distortion of the pulse during its propagation, as the group velocity dispersion vanishes at low frequencies. Conventional wisdom tells us that the group velocity, which is the velocity of the peak of the pulse, is the true signal velocity, in contradistinction to the phase velocity, since normally energy transport is characterized by the group and not the phase velocity. If we were to cling to this definition of signal velocity, then we would be forced to accept signal velocities faster than light. However, special relativity is in fact not violated by these superluminal group velocities, as we shall see in the next section.

Unlike a medium in its ground state, the inverted medium can temporarily loan part of its stored energy to the forward tail of the wave packet, in a pulse-reshaping process which moves the peak of the wave packet forward in time. One can think of this pulse-reshaping process as the virtual amplification of the forward tail of the wave packet, followed by the virtual absorption of the peak, resulting in an *advancement* of the wave packet. This is a reversal of the pulse-reshaping process produced by the uninverted medium, in which the peak of a wave packet first undergoes virtual absorption, followed by the virtual amplification of its trailing tail, resulting in a *retardation* of the wave packet. Energy is loaned by the medium to the wave, or vice versa, in the inverted and the uninverted cases, respectively, so that the energy in the pulse remains unchanged in both kinds of pulse-reshaping processes in these *transparent* media. Thus the energy velocity, as defined by Sommerfeld and Brillouin (Brillouin [1960]), is also superluminal for the inverted medium near zero frequency

$$\upsilon_{\mathrm{E}}(0) \equiv \frac{\langle S\rangle}{\langle u\rangle} = \frac{c}{\sqrt{\varepsilon(0)}} = \frac{c}{n(0)} = \upsilon_{\mathrm{p}}(0) > c\,, \tag{7.8}$$

where $\langle S\rangle$ is the time-averaged Poynting vector, $\langle u\rangle$ is the time-averaged energy density, and $\varepsilon(0)$ is the zero-frequency dielectric constant. This is a reversal of the case of the uninverted medium, where the energy velocity is of course subluminal. The Sommerfeld and Brillouin energy velocity is usually interpreted as the velocity of energy transport by the propagating wave packet. However, there is controversy concerning the proper definition of the energy velocity (Schulz-DuBois [1969], Loudon [1970], Oughstun and Shen [1988], Diener [1997]); after all, in addition to the purely electromagnetic energy density, there is energy stored in the inverted medium itself.

Still more surprisingly, the "signal" velocity of Sommerfeld and Brillouin, which they defined *arbitrarily* as the propagation velocity of the first point of half-maximum wave amplitude, is the same here as the group velocity, since there is little distortion of the shape of the wave packet during its propagation. However, we shall see that it is highly misleading to call this the "signal" velocity. Since dispersion is negligible in this large, transparent spectral window stretching from DC to the low-frequency side of resonance, all of the above wave velocities, including the so-called "signal" velocity, are faster than c. It should be emphasized that any arbitrary, low-frequency finite-bandwidth wave form, e.g., Rachmaninov's 3rd Piano Concerto, and not merely Gaussian wave packets, will propagate faster than *c with negligible distortion*, so that a complicated wave form can also be advanced to earlier times at the output face of the inverted medium.

Recently, some of these counterintuitive effects have been observed in an experiment with very low frequency bandpass electronic amplifiers (Mitchell and Chiao [1997]). Negative group delays were observed, in which pulses transmitted through a chain of amplifiers were *advanced* with little distortion by several milliseconds, i.e., the transmitted peak left the output port of the amplifier chain *before* the incident peak arrived at the input port. However, the behavior of abrupt "fronts" and "backs" showed that causality was in fact not violated.

The nervous reader may ask at this point how it is possible to avoid a violation of special relativity. A brief answer is that the *front* velocity of Sommerfeld and Brillouin in the case of a medium with inverted populations is still exactly c, as it is also in the case of tunneling. This will be shown in detail in the next section. We shall further see that the front velocity, and not the so-called "signal" velocity of Sommerfeld and Brillouin, should be identified as the true signal velocity, and this fact will prohibit any genuine information from being communicated faster than c. The reader may also object to our use of the Lorentz model, which after all is merely a model. However, the above results can also be shown to follow very generally from the Kramers–Kronig relations, which are themselves consequences of causality and linearity. These results must therefore transcend all models (Chiao [1993]. In general, the Kramers–Kronig relations (i.e., the very

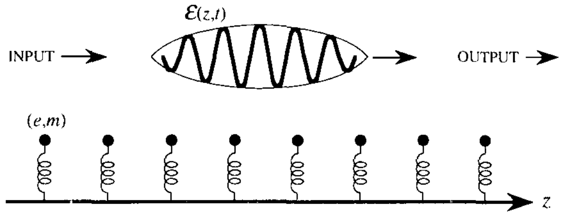

Fig. 10. A linear array of undamped Lorentz oscillators for calculating the polariton-like (for uninverted atoms) and the tachyon-like dispersion relations (for inverted atoms), for a strongly coupled atom-radiation system.

requirement of causality itself) demand that superluminal group velocities arise in *any* dispersive medium (Bolda, Chiao and Garrison [1993]); in particular, they must arise in any medium with gain.

As an example of the second, *transient* situation, we shall focus on the special case of tachyon-like propagation of wave packets through a population-inverted medium at frequencies close to resonance. Although the theory for the tachyon-like excitations of this medium was originally worked out starting from the sine-Gordon equation for the fully nonlinear problem of the coupling between the two-level atoms and the radiation field (Chiao, Kozhekin and Kurizki [1996]), we present here a simplified, linearized version of this theory, which brings out more directly the essential features. Our goal is to calculate the dispersion relations for small-amplitude excitations of the strongly coupled atom-field medium, and show that tachyon-like excitations emerge naturally as the normal modes of an undamped medium composed of atoms with suddenly inverted populations.

Consider a long collection of Lorentz oscillators with a uniform density along the z-axis (see fig. 10). (There are no mirrors at the ends of this medium.) We shall focus on the special case of *undamped* motions of these oscillators. Such a system is a good model for two-level atoms in their ground states (Burnham and Chiao [1969]), but can be generalized easily to the case of atoms with inverted populations (see eq. 7.2). The two equations which describe the coupled atom-radiation system are (i) Maxwell's equations in the form of the wave equation

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = \frac{4\pi}{c^2}\frac{\partial^2 P}{\partial t^2}, \tag{7.9}$$

and (ii) the undamped simple harmonic equation of motion for the Lorentz oscillators,

$$\frac{\partial^2 x}{\partial t^2} + \omega_0^2 x = \frac{eE}{m}. \tag{7.10}$$

Here $P = Nex$ is the polarization of the medium (N being the number density of Lorentz oscillators and x being the displacement from equilibrium of a given oscillator), and ω_0 is the natural resonance frequency of the oscillators.

In order to calculate how a wave packet will propagate through this system, we shall use the slowly-varying envelope ansatz (SVEA)

$$\begin{aligned} E &= \mathcal{E}(z,t)\exp[\mathrm{i}(k_0 z - \omega_0 t)], \\ P &= \mathcal{P}(z,t)\exp[\mathrm{i}(k_0 z - \omega_0 t)], \\ x &= \chi(z,t)\exp[\mathrm{i}(k_0 z - \omega_0 t)], \end{aligned} \tag{7.11}$$

where $\mathcal{E}(z,t)$, $\mathcal{P}(z,t)$, $\chi(z,t)$ are all slowly varying envelopes which modulate the common, fast plane-wave factor, $\exp[\mathrm{i}(k_0 z - \omega_0 t)]$, and where by definition, $k_0 = \omega_0/c$ is the vacuum wave number of the uncoupled waves. Neglecting the second derivatives of the slowly-varying amplitudes, we obtain two first-order partial differential equations (PDE's):

$$2\mathrm{i}k_0\frac{\partial\mathcal{E}}{\partial z} + 2\mathrm{i}\frac{\omega_0}{c^2}\frac{\partial\mathcal{E}}{\partial t} = -\frac{4\pi\omega_0{}^2 Ne}{c^2}\chi \tag{7.12}$$

$$-2\mathrm{i}\omega_0\frac{\partial\chi}{\partial t} = \frac{e\mathcal{E}}{m} \tag{7.13}$$

which are the linearized Maxwell–Bloch equations. Taking the partial derivative with respect to time of the first of these equations, and eliminating $\partial\chi/\partial t$ by means of the second equation, we obtain a PDE for the electric field envelope:

$$\frac{\partial^2\mathcal{E}}{\partial z\partial t} + \frac{1}{c}\frac{\partial^2\mathcal{E}}{\partial t^2} + \frac{1}{4}\frac{\omega_\mathrm{p}^2}{c}\mathcal{E} = 0\,. \tag{7.14}$$

To include the possibility of population inversion, we use the effective plasma frequency ω_p given by eq. (7.2)[12]. In order to find the dispersion relations, we substitute into this PDE the plane-wave ansatz

$$\mathcal{E} = A\exp[\mathrm{i}(Kz - \Omega t)]\,, \tag{7.15}$$

where $K \equiv k - k_0$ and $\Omega \equiv \omega - \omega_0$; this converts eq. (7.14) into the algebraic (quadratic) equation

$$\Omega^2 - Kc\Omega - \tfrac{1}{4}\omega_\mathrm{p}{}^2 = 0\,. \tag{7.16}$$

The solution of this quadratic equation yields the dispersion relations

$$\Omega = \tfrac{1}{2}Kc \pm \tfrac{1}{2}\left(K^2c^2 + \omega_\mathrm{p}{}^2\right)^{1/2}, \tag{7.17}$$

which are plotted in fig. 11. In the case of uninverted atoms ($w = -1$), ω_p is real, and we recover polariton-like dispersion relations, whereas in the case of inverted

[12] The definition of the effective plasma frequency used here differs from that used in Chiao, Kozhekin and Kurizki [1996] in that the factor of $(-w)$ there has been absorbed into $\omega_\mathrm{p}{}^2$ here.

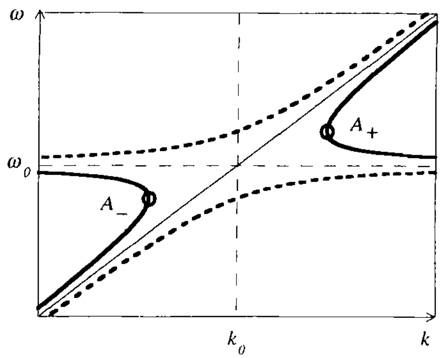

Fig. 11. Dispersion relations for the coupled atom-radiation system as calculated from the undamped Lorentz model for uninverted atoms (dashed curve), corresponding to polaritonic branches with $w = -1$, and for inverted atoms (solid curve), corresponding to tachyonic branches with $w = +1$.

atoms ($w = +1$), ω_p is imaginary, and we find tachyon-like dispersion relations. The tachyonic branches have group velocities which are always faster than c (but which approach c far from resonance), infinite at the turning points A_+ and A_-, or negative as resonance is approached. Computer simulations indicate that negative group velocities in gain media also have a well-defined physical meaning (Bolda, Garrison and Chiao [1994]). However, under no circumstances can these tachyonic excitations outrace the front (Aharonov, Komar and Susskind [1969], Chiao, Kozhekin and Kurizki [1996]).

The wave-number gap between A_+ and A_- is a gap of instability arising from population inversion. Vacuum fluctuations with frequency components inside this gap can trigger spontaneous emission, and hence superfluorescence. However, spontaneous emission does not prevent superluminality. It has been shown that the typical delay time for the onset of superfluorescence in realistic media (Bolda [1996]) is much longer than the passage time for a typical tachyonic excitation, so that the population inversion does not disappear due to the emission of a superfluorescent pulse before the tachyonic excitation has had a chance to finish propagating through the medium. This should make experiments to observe tachyon-like excitations possible, and an experiment has been commenced at Berkeley to demonstrate the existence of these excitations in ammonia gas pumped by a carbon dioxide laser, on the same transition used in the first maser by Gordon, Zeiger and Townes [1954].

It has also been shown that the effective plasma frequency is directly proportional to the effective mass of the corresponding collective excitation; hence a polariton-like excitation possesses a *real* effective mass, but a tachyon-like excitation possesses an effective mass which is *imaginary*, which is the basis for calling them "tachyonic" (Chiao, Kozhekin and Kurizki [1996]). However, it should be emphasized that these tachyonic excitations should be viewed as quasi-particles *in a medium*, like phonons, and not as true particles *in the vacuum*, like photons.

§ 8. Why Is Einstein Causality Not Violated?

The question naturally arises whether Einstein causality is or is not violated by the superluminal behavior exhibited in tunneling or in population-inverted media. In the case of tunneling, numerous theoretical analyses have shown that there is in fact no contradiction with causality (see, for example, Deutch and Low [1993], Hass and Busch [1994], Azbel [1994], Wang and Zhang [1995] and Japha and Kurizki [1996b]). Let us first make some qualitative remarks concerning this question, and then return to some more rigorous, quantitative considerations. We shall restrict our attention here to classical electromagnetic signals, for example, voltage wave forms displayed on an oscilloscope. Also, we shall assume the total absence of noise in the following section. However, the fundamental considerations of causality given below for classical electromagnetism should be generalizable to quantum field theories (Eberhard and Ross [1988]).

The qualitative discussion starts with the observation that there is no information contained in the peak of an analytic wave packet which is not already present in its forward tail. For example, the behavior near a peak of an analytic wave form, e.g., of a Gaussian wave packet, could have been predicted by Taylor's theorem from the earlier behavior of its forward exponential tail (i.e., using the knowledge of all the derivatives of the earlier portions of the wave form, we could extrapolate to all later portions; in particular, we could in principle predict the exact moment of arrival for the peak of the wave form [13]). Therefore there is no real surprise when the peak eventually arrives. New information is communicated only when there is an *unexpected* change, such as a discontinuity, whose arrival time *cannot* be inferred from the past behavior of the wave.

[13] *Pulse reshaping* mechanisms, such as the virtual amplification of the forward tail followed by the virtual absorption of the peak of the Gaussian wave packet which reproduces the shape of this wave packet, can therefore advance the peak forward in time in a completely predictable and causal manner.

A simple example of such a discontinuity is that of a step-modulated sine wave, i.e., a jump discontinuity or "front", which Sommerfeld and Brillouin used in their study of precursors. Their wave form thus has a sharp jump from zero to finite intensity at the front. They found that no features of their solution, including their precursors, could ever overtake this front. In contrast to the peak of the Gaussian wave packet, the arrival of the front could never have been predicted from any prior information, and hence the front in this example constitutes a genuine signal, i.e., new information.

However, any point of nonanalyticity in a wave form, such as a jump discontinuity in some higher derivative, and not just a jump discontinuity in the wave amplitude such as the front of Sommerfeld and Brillouin, can serve as a carrier of genuinely new information. Any such point of nonanalyticity is always preserved upon transmission by any linear, causal system, as we shall demonstrate below. *Nonanalytic* wave forms, for example, piecewise analytic functions joined smoothly at given points of nonanalyticity, have Fourier components which fall off algebraically in the high-frequency limit (the higher the order of the derivative jump, the larger the negative exponent of the frequency in this fall-off). It is the response in the *infinite-frequency* limit of the system that ultimately determines the propagation speed of the points of nonanalyticity, and hence of truly new information. Since the propagation of infinite-frequency components of a disturbance occurs at the vacuum speed of light, i.e., at Sommerfeld and Brillouin's front velocity, this is also the velocity of propagation of the points of nonanalyticity, and hence of genuine information. It is fundamentally for this reason that Einstein causality cannot be violated under any circumstances, either in the tunneling barrier or in population-inverted media.

The rigorous, quantitative considerations start with a "black box" which locally relates an input to an output wave form by means of a *linear* transfer function $T(\tau)$, via the equation

$$f_{\text{out}}(t) = \int_{-\infty}^{\infty} T(\tau) f_{\text{in}}(t-\tau)\, \mathrm{d}\tau\,, \tag{8.1}$$

where τ is a delay time, $f_{\text{in}}(t)$ is an arbitrary input function, and $f_{\text{out}}(t)$ is the resulting output function. For example, the input $f_{\text{in}}(t)$ could represent an electric field applied to an atom, whose polarizability would be represented by $T(\tau)$, and the output $f_{\text{out}}(t)$ would represent the dipole moment response of the atom produced by the electric field. It should be stressed that $f_{\text{in}}(t)$ and $f_{\text{out}}(t)$ can represent any of the higher derivatives of the wave form, as well as the wave form itself. This follows directly from the *linearity* of eq. (8.1).

The principle of causality demands that the integrand must vanish for $\tau < 0$ in eq. (8.1), since any effect (e.g., the atomic dipole moment) *must not precede* its cause (e.g., the applied electric field). This necessitates that

$$T(\tau) = 0 \text{ for all } \tau < 0 . \tag{8.2}$$

When eq. (8.1) is Fourier tranformed into the frequency domain, it becomes

$$\widetilde{f_{\text{out}}}(\omega) = \widetilde{T}(\omega)\widetilde{f_{\text{in}}}(\omega) , \tag{8.3}$$

where the tildes denote Fourier transforms. The complex frequency transfer function $\widetilde{T}(\omega)$, as a consequence of eq. (8.2), must satisfy the condition that

$$\widetilde{T}(\omega) \text{ is analytic for all } \operatorname{Im} \omega > 0 , \tag{8.4}$$

i.e., the complex frequency transfer function must be analytic in the upper half frequency plane (UHP), which is an expression of causality equivalent to eq. (8.2). This leads to the Kramers–Kronig relations for $\widetilde{T}(\omega)$ (Landau and Lifshitz [1960]).

Now suppose that the function $f_{\text{in}}(t)$ has a front in it at the time t_0, so that

$$f_{\text{in}}(t) = 0 \text{ for all } t < t_0 . \tag{8.5}$$

Then the Fourier transform of this function must satisfy the condition that

$$\widetilde{f_{\text{in}}}(\omega) \text{ is analytic for all } \operatorname{Im} \omega > 0 , \tag{8.6}$$

i.e., the Fourier transform of the input function must be analytic in the UHP. Since each of its factors are analytic in the UHP, it follows that the product

$$\widetilde{f_{\text{out}}}(\omega) = \widetilde{T}(\omega)\widetilde{f_{\text{in}}}(\omega) \text{ is analytic for all } \operatorname{Im} \omega > 0 , \tag{8.7}$$

i.e., the Fourier transform of the output function must also be analytic in the UHP. Therefore using the inverse Fourier transform, we obtain the result

$$f_{\text{out}}(t) = 0 \text{ for all } t < t_0' , \tag{8.8}$$

where it can be shown that $t_0' = t_0$ for any "black box" that has a negligible spatial extent. This proves that fronts in the input survive the transfer through any "black box" which is linear and causal: *Fronts are preserved in the output.* Therefore, although there is no physical law which guarantees that an incoming peak turns into an outgoing peak, there *is* a physical law namely causality, that guarantees that an incoming front turns into an outgoing front, even when the front carries little energy or probability.

Using linearity, we can generalize this result to any point of nonanalyticity, for example, a jump discontinuity in some higher derivative of the wave form. Using the superposition principle, which also follows from the linearity of the system, we can further generalize this to all the points of nonanalyticity $t_0, t_1, t_2, \ldots$ in the wave form. Motivated by these considerations, we shall define a signal as the complete set of all the points of nonanalyticity $\{t_0, t_1, t_2, \ldots\}$, together with the values of the input function $f_{\rm in}(t)$ in a small but finite interval of time inside the domain of analyticity immediately following these points. It should be emphasized that this definition leads to a signal velocity that differs from the conventional one given by the group velocity. The principle of causality makes this new definition necessary. However, we are making idealizations, in particular, in assuming the highest possible detector sensitivity and the perfect noiselessness of the system, in formulating this *fundamental* definition, but this may not be a *practical* definition under all circumstances.

The generalization of this argument to *propagation* through any *spatially extended* "black box" that is linear and causal, is straightforward (Jackson [1975]). For an input with a single point of nonanalyticity at t_0 given by

$$f_{\rm in}(t) = 0 \text{ for all } t < t_0\,, \tag{8.9}$$

the output must satisfy the condition that

$$f_{\rm out}(t) = 0 \text{ for all } t - d/c < t_0\,, \tag{8.10}$$

where d is the distance from the input face to the output face of the "black box". Using the definition given above, we conclude that genuine signals cannot propagate faster than c. In fact they propagate exactly at c, i.e., at the front velocity. Thus Einstein causality, i.e., special relativity, is not violated.

Although at a fundamental level no genuine signal can be transmitted faster than light, at a practical level there are situations in which useful temporal advances of a wave form are possible. For example, unwanted positive group delays arising from normal dielectric media in the system may be compensated by negative group delays, but only up to the limit permitted by Einstein causality (Chiao, Boyce and Garrison [1995], Steinberg and Chiao [1994b]). In another example, a detector followed by a discriminator with a fixed trigger level can register the arrival of a pulse earlier with the aid of an amplifier than without it, but again only up to the Einsteinian limit (Chiao [1996], Mitchell and Chiao [1997]).

The meaning of superluminal group velocities was also considered recently by Diener [1996]. He also concluded that *superluminal* group velocities cannot be interpreted as a velocity of information transfer. The method he used to reach this conclusion was different, being based on the Green's function and its application to the analytic continuation of the pulse shape using information only within the past light cone. However, Diener continued to interpret *subluminal* group velocities as signal velocities, whereas we believe that the *same* definition for "signal" should in principle be consistently applied to *both* superluminal *and* subluminal cases.

§ 9. Conclusion

We thus see that a relatively old debate over how long the tunneling process takes has begun to shed new light on a variety of issues, in no small part thanks to the realization that the analogy between electromagnetic and Schrödinger wave equations permits the same phenomenon to be studied in optics rather than in the solid state. We are developing a new understanding of the limits imposed by causality on various propagation speeds, and have relearned that a group velocity, and even the motion of a real, well-behaved wave packet peak, can in fact be greater than c. We see also that time in quantum mechanics is not a simple issue: a given process may have not a single duration, but a set of different timescales describing its various aspects. When the problem is studied in the light of particle–wave duality, where the actual time of arrival of individual quanta is on average earlier than what would be expected from a naïve application of causality principles, we come up against one of the central problems of quantum mechanics – the extent to which one can discuss quantities which have not been measured directly, such as the past history of a particle we observe at the present time. This applies to single-photon wave packet propagation both in tunneling and in gain media. In the case of tunneling, there is no clear way to separate "to-be-transmitted" and "to-be-reflected" portions, nor to answer the question of *where* a particle is save in a probabilistic manner. Yet a quantum particle may be forced eventually into a purely transmitted or reflected state, and the question of how much effect it has had on devices placed in its path (or how much effect they have had on it) is certainly a reasonable, and an important, one to ask.

The superluminality of the tunneling process should also be a relevant consideration in fundamental questions concerning the nature of Hawking radiation from an evaporated black hole, and of similar radiative processes which involve the tunneling of particles through an event horizon (Massar and Parentani [1997]). Closely related are the questions raised here: what constitutes a signal, i.e., what is information at the quantum level? Aside from their fundamental interest, the answers to such questions are crucial for responding to questions such as what the maximum speed of a tunneling device might be. Work continues on these issues at both the experimental and the theoretical level, and in both arenas, optical versions of tunneling and other superluminal phenomena have been and will continue to be of great value to the debate. Not only should we expect this work to teach us more about the fundamental nature of the tunneling process, but about some of the deepest mysteries of quantum mechanics.

Acknowledgments

We thank Jack Boyce, John Garrison, Eivind Hauge, Rolf Landauer, John McGuire, and Morgan Mitchell for helpful discussions, and especially Jack Boyce for his help in producing the manuscript. This work was supported by the ONR under Grant No. N000149610034.

Note added in proof

After this review was written, another review on a similar subject was published by Nimtz and Heitmann [1997] (Prog. Quantum Electron. **21**, 81). These authors deny the central significance of the front velocity for signals. For the reasons given in § 8, we believe that their point of view is fundamentally incorrect.

References

Aharonov, Y., D.Z. Albert and L. Vaidman, 1988, Phys. Rev. Lett. **60**, 1351.

Aharonov, Y., A. Komar and L. Susskind, 1969, Phys. Rev. **182**, 1400.

Aharonov, Y., and L. Vaidman, 1990, Phys. Rev. A **41**, 11.

Anderson, M.H., J.R. Ensher, M.R. Matthews, C.E. Wieman and E.A. Cornell, 1995, Science **269**, 198.

Andrews, M.R., C.G. Townsend, H.-J. Miesner, D.S. Durfee, D.M. Kurn and W. Ketterle, 1997, Science **275**, 637.

Ashcroft, N.W., and N.D. Mermin, 1976, Solid State Physics (Holt, Rinehart and Winston, New York) p. 146.

Azbel, M.Ya., 1994, Solid State Commun. **91**, 439.

Balcou, Ph., and L. Dutriaux, 1997, Phys. Rev. Lett. **78**, 851.

Baz', A.I., 1967, Sov. J. Nucl. Phys. **5**, 161.
Bohm, D., 1951, Quantum Theory (Prentice Hall, Englewood Cliffs, NJ) p. 290.
Bohm, D., 1952, Phys. Rev. **85**, 166.
Bolda, E.L., 1996, Phys. Rev. A **54**, 3514.
Bolda, E.L., R.Y. Chiao and J.C. Garrison, 1993, Phys. Rev. A **48**, 3890.
Bolda, E.L., J.C. Garrison and R.Y. Chiao, 1994, Phys. Rev. A **49**, 2938.
Born, M., and E. Wolf, 1975, Principles of Optics (Pergamon Press, Oxford).
Boyd, R.W., 1992, Nonlinear Optics (Academic Press, San Diego).
Brillouin, L., 1960, Wave Propagation and Group Velocity (Academic Press, New York).
Brodowsky, H.M., W. Heitmann and G. Nimtz, 1996, Phys. Lett. A **222**, 125.
Burnham, D.C., and R.Y. Chiao, 1969, Phys. Rev. **188**, 667.
Büttiker, M., 1983, Phys. Rev. B **27**, 6178.
Büttiker, M., 1990, in: Electronic Properties of Multilayers and Low Dimensional Semiconductors, eds J.M. Chamberlain, L. Eaves and J.C. Portal (Plenum Press, New York).
Büttiker, M., and R. Landauer, 1982, Phys. Rev. Lett. **49**, 1739.
Büttiker, M., and R. Landauer, 1985, Phys. Scr. **32**, 429.
Carniglia, C.K., and L. Mandel, 1971a, J. Opt. Soc. Am. **61**, 1035.
Carniglia, C.K., and L. Mandel, 1971b, J. Opt. Soc. Am. **61**, 1423.
Chiao, R.Y., 1993, Phys. Rev. A. **48**, R34.
Chiao, R.Y., 1994, Quantum Opt. **6**, 359.
Chiao, R.Y., 1996, Population inversion and superluminality, in: Amazing Light: A Volume Dedicated to Charles Hard Townes on his 80th Birthday, ed. R.Y. Chiao (Springer, New York) p. 91.
Chiao, R.Y., E.L. Bolda, J. Bowie, J. Boyce, J.C. Garrison and M.W. Mitchell, 1995, Quantum and Semiclassical Opt. **7**, 279.
Chiao, R.Y., and J. Boyce, 1994, Phys. Rev. Lett. **73**, 3383.
Chiao, R.Y., J. Boyce and J.C. Garrison, 1995, in: Fundamental Problems in Quantum Theory, eds D.M. Greenberger and A. Zeilinger, Ann. N.Y. Acad. Sci. **755**, 400.
Chiao, R.Y., A.E. Kozhekin and G. Kurizki, 1996, Phys. Rev. Lett. **77**, 1254.
Chiao, R.Y., P.G. Kwiat and A.M. Steinberg, 1991, Physica B **175**, 257.
Chiao, R.Y., P.G. Kwiat and A.M. Steinberg, 1993, Sci. Am. **269**, 52.
Chu, S., and S. Wong, 1982, Phys. Rev. Lett. **48**, 738; **49**, 1293.
Condon, E.U., 1931, Rev. Mod. Phys. **3**, 43; see pp. 74ff.
Deutch, J.M., and F.E. Low, 1993, Ann. Phys. **228**, 184.
Deutsch, I.H., 1991, Am. J. Phys. **59**, 834.
Deutsch, I.H., and J.C. Garrison, 1991, Phys. Rev. A **43**, 2498.
Deutsch, M., and J.E. Golub, 1996, Phys. Rev. A 53, 434.
Dewdney, C., L. Hardy and E.J. Squires, 1993, Phys. Lett. A **184**, 6.
Dewdney, C., and B.J. Hiley, 1982, Found. Phys. **12**, 27.
Diener, G., 1996, Phys. Lett. A **223**, 327.
Diener, G., 1997, Energy transport in dispersive media and superluminal group velocities, preprint.
Dürr, D., W. Fusseder, S. Goldstein and N. Zanghi, 1993, Z. Naturforschung **48A**, 1261.
Eberhard, P.H., and R.R. Ross, 1988, Found. Phys. Lett. **2**, 127.
Eisenbud, L., 1948, Ph.D. Thesis (Princeton University).
Enders, A., and G. Nimtz, 1992, J. Phys. I (France) **2**, 1693.
Enders, A., and G. Nimtz, 1993, J. Phys. I (France) **3**, 1089.
Englert, B.-G., M.O. Scully, G. Sussmann and H. Walther, 1992, Z. Naturforschung **47A**, 1175.
Esteve, D., J.M. Martinis, C. Urbina, E. Turlot and M.H. Devoret, 1989, Phys. Scripta T **29**, 121.
Falck, J.P., and E.H. Hauge, 1988, Phys. Rev. B **38**, 3287.
Fertig, H.A., 1990, Phys. Rev. Lett. **65**, 2321.

Fertig, H.A., 1993, Phys. Rev. B **47**, 1346.
Garrett, C.G.B., and D.E. McCumber, 1970, Phys. Rev. A **1**, 305.
Gasparian, V., M. Ortuño, J. Ruiz, E. Cuevas and M. Pollak, 1995, Phys. Rev. B **51**, 6743.
Gasparian, V., and M. Pollak, 1993, Phys. Rev. B **47**, 2038.
Gaylord, T.K., G.N. Henderson and E.N. Glytsis, 1993, J. Opt. Soc. Am. B **10**, 333.
Ghatak, A.K., and S. Banerjee, 1989, Appl. Opt. **28**, 1960.
Ghatak, A.K., M.R. Shenoy, I.C. Goyal and K. Thyagarajan, 1986, Opt. Commun. **56**, 313.
Giakos, G.C., and T.K. Ishii, 1991a, Microwave and Opt. Technol. Lett. **1**, 374.
Giakos, G.C., and T.K. Ishii, 1991b, Microwave and Opt. Technol. Lett. **4**, 128.
Giakos, G.C., and T.K. Ishii, 1991c, Microwave and Opt. Technol. Lett. **4**, 79.
Glauber, R.J., 1965, in: Quantum Optics and Electronics, Ecole d'été de physique théorique, Les Houches, Haute-Savoie, France, eds C. DeWitt, A. Blandin and C. Cohen-Tannoudji (Gordon and Breach, New York) p. 63.
Gordon, J.P., H. Zeiger and C.H. Townes, 1954, Phys. Rev. **95**, 282L.
Guéret, P., A. Baratoff and E. Marclay, 1987, Europhys. Lett. **3**, 367.
Guéret, P., E. Marclay and H. Meier, 1988a, Appl. Phys. Lett. **53**, 1617.
Guéret, P., E. Marclay and H. Meier, 1988b, Solid State Commun. **68**, 977.
Hänggi, P., 1993, Path integral approach to interaction and tunneling times, in: Lectures on Path Integration, eds H.A. Cerdeira, L. Lundqvist, D. Mugnai, A. Ranfagni, V. Sayakanit and L. Schulman (World Scientific, Singapore).
Hartman, T.E., 1962, J. Appl. Phys. **33**, 3427.
Hass, K., and P. Busch, 1994, Phys. Lett. A **185**, 9.
Hauge, E.H., J.P. Falck and T.A. Fjeldly, 1987, Phys. Rev. B **36**, 4203.
Hauge, E.H., and J.A. Støvneng, 1989, Rev. Mod. Phys. **61**, 917.
Heitmann, W., and G. Nimtz, 1994, Phys. Lett. A **196**, 154.
Hino, M., N. Achiwa, S. Tasaki, T. Ebisawa, T. Akiyoshi and T. Kawai, 1996, Traversal time through magnetic thin film using Larmor precession, in: Proc. Neutron Optics Kumatori '96, J. Phys. Soc. Japan, Suppl. A **65**, to appear.
Holland, P.R., 1993, Phys. Rep. **224**, 95.
Hong, C.K., Z.Y. Ou and L. Mandel, 1987, Phys. Rev. Lett. **59**, 2044.
Iannaccone, G., 1996, Weak measurement and the traversal time problem, Adriatico Research Conference on Tunnelling and its implications, ICTP, Trieste: preprint quant-ph/9611018.
Ishii, T.K., and G.C. Giakos, 1991, Microwaves and RF **30**, 114 and references therein.
Jackson, J.D., 1975, Classical Electrodynamics (Wiley, New York).
Japha, Y., and G. Kurizki, 1996a, Phys. Rev. Lett. **11**, 2909.
Japha, Y., and G. Kurizki, 1996b, Phys. Rev. A **53**, 586.
Jeffers, J., and S.M. Barnett, 1993, Phys. Rev. A **47**, 3291.
John, S., 1991, Phys. Today **44**, 32.
Kittel, C., 1986, Introduction to Solid State Physics (Wiley, New York) p. 372.
Krenzlin, H.M., J. Budezies and K.W. Kehr, 1996, Phys. Rev. A **53**, 3749.
Kurizki, G., 1997, in preparation.
Kurizki, G., and Y. Japha, 1993, in: Proc. Adriatico Research Workshop on Quantum Interferometry (Trieste), organized by F. de Martini and A. Zeilinger.
Landau, L.D., and E.M. Lifshitz, 1960, Electrodynamics of Continuous Media (Pergamon Press, Oxford) p. 256.
Landauer, R., 1989, Nature **341**, 567.
Landauer, R., and Th. Martin, 1994, Rev. Mod. Phys. **66**, 217.
Leavens, C.R., 1990, Solid State Commun. **74**, 923.
Leavens, C.R., 1993, Phys. Lett. A **178**, 27.

Leavens, C.R., 1995, Phys. Lett. A **197**, 88.
Leavens, C.R., and G.C. Aers, 1989, Phys. Rev. B **40**, 5387.
Leavens, C.R., and G.C. Aers, 1991, Solid State Commun. **78**, 1015.
Leavens, C.R., and G.C. Aers, 1993, in: Scanning Tunneling Microscopy III, eds R. Wiesendanger and H.J. Güntherodt (Springer, Berlin).
Leavens, C.R., G. Iannaccone and W.R. McKinnon, 1995, Phys. Lett. A **208**, 17.
Leavens, C.R., and W.R. McKinnon, 1995, Phys. Rev. A **51**, 2748.
Lee, B., 1996, in: O.S.A. Annual Meeting Abstracts, p. 185.
Lee, B., and W. Lee, 1995, Superlattices and Microstructures **18**, 277.
Lee, B., and W. Lee, 1997, J. Opt. Soc. Am. B **14**, 777.
Loudon, R., 1970, J. Phys. A **3**, 233.
Low, F.E., and P.F. Mende, 1991, Ann. Phys. **210**, 380.
MacColl, L.A., 1932, Phys. Rev. **40**, 621.
Martin, Th., and R. Landauer, 1992, Phys. Rev. A **45**, 2611.
Massar, S., and R. Parentani, 1997, Phys. Rev. Lett. **78**, 3810.
Mewes, M.-O., M.R. Andrews, D.M. Kurn, D.S. Durfee, C.G. Townsend and W. Ketterle, 1997, Phys. Rev. Lett. **78**, 582.
Mitchell, M.W., and R.Y. Chiao, 1997, Phys. Lett. A **230**, 133.
Muga, J.G., S. Brouard and R. Sala, 1992a, Phys. Lett. A **167**, 24.
Muga, J.G., S. Brouard and R. Sala, 1992b, J. Phys.: Condens. Matter (UK) **4**, L579.
Mugnai, D., A. Ranfagni, R. Ruggeri and A. Agresti, 1992, Phys. Rev. Lett. **68**, 259.
Mugnai, D., A. Ranfagni, R. Ruggeri and A. Agresti, 1994, Phys. Rev. E **49**, 1771.
Mugnai, D., A. Ranfagni and L.S. Schulman, 1997, Phys. Rev. E **55**, 3593.
Nimtz, G., and W. Heitmann, 1997, Prog. Quantum Electron. **21**, 81.
Nimtz, G., H. Spieker and H.M. Brodowsky, 1994, J. Phys. I (France) **4**, 1379.
Niu, Q., X.-G. Zhao, G.A. Georgakis and M.G. Raizen, 1996, Phys. Rev. Lett. **76**, 4504.
Olkhovsky, V.S., and E. Recami, 1992, Phys. Rep. **214**, 339.
Oughstun, K.E., and S. Shen, 1988, J. Opt. Soc. Am. B **5**, 2395.
Pollak, E., and W.H. Miller, 1984, Phys. Rev. Lett. **53**, 115.
Raciti, F., and G. Salesi, 1994, J. Phys. I (France) **4**, 1783.
Ranfagni, A., P. Fabeni, G.P. Pazzi and D. Mugnai, 1993, Phys. Rev. E **48**, 1453.
Ranfagni, A., D. Mugnai, P. Fabeni and G.P. Pazzi, 1990, Phys. Scripta **42**, 508.
Ranfagni, A., D. Mugnai, P. Fabeni and G.P. Pazzi, 1991, Appl. Phys. Lett. **58**, 774.
Ranfagni, A., D. Mugnai, P. Fabeni, G.P. Pazzi, G. Naletto and C. Sozzi, 1991, Physica B **175**, 283.
Rybachenko, V.F., 1967, Sov. J. Nucl. Phys. **5**, 635.
Schulz-DuBois, E.O., 1969, Proc. IEEE **57**, 1748.
Segard, B., and B. Macke, 1985, Phys. Lett. A **109**, 213.
Shapiro, J.H., and K.X. Sun, 1994, J. Opt. Soc. Am. B **11**, 1130.
Sokolovski, D., 1995, Phys. Rev. A **52**, R5.
Sokolovski, D., and L.M. Baskin, 1987, Phys. Rev. A **36**, 4604.
Sokolovski, D., and J.N.L. Connor, 1990, Phys. Rev. A **42**, 6512.
Sokolovski, D., and J.N.L. Connor, 1993, Phys. Rev. A **47**, 4677.
Sokolovski, D., and J.N.L. Connor, 1994, Solid State Commun. **89**, 475.
Sommerfeld, A., 1907, Phys. Z. **8**, 841.
Spector, J., H.L. Stormer, K.W. Baldwin, L.N. Pfeiffer and K.W. West, 1990, Appl. Phys. Lett. **56**, 2433.
Spielmann, Ch., R. Szipöcs, A. Stingl and F. Krausz, 1994, Phys. Rev. Lett. **73**, 2308.
Steinberg, A.M., 1994, J. Phys. I (France) **4**, 1813.
Steinberg, A.M., 1995a, Phys. Rev. A **52**, 32.

Steinberg, A.M., 1995b, Phys. Rev. Lett. **74**, 2405.
Steinberg, A.M., 1995c, La Recherche **281**, 46.
Steinberg, A.M., and R.Y. Chiao, 1994a, Phys. Rev. A **49**, 3283.
Steinberg, A.M., and R.Y. Chiao, 1994b, Phys. Rev. A **49**, 2071.
Steinberg, A.M., and R.Y. Chiao, 1995, Phys. Rev. **51**, 3525.
Steinberg, A.M., P.G. Kwiat and R.Y. Chiao, 1992a, Phys. Rev. Lett. **68**, 2421.
Steinberg, A.M., P.G. Kwiat and R.Y. Chiao, 1992b, Phys. Rev. A **45**, 6659.
Steinberg, A.M., P.G. Kwiat and R.Y. Chiao, 1993, Phys. Rev. Lett. **71**, 708.
Steinberg, A.M., P.G. Kwiat and R.Y. Chiao, 1994, Found. Phys. Lett. **7**, 223.
Steinberg, A.M., R. Thompson, V. Bagnoud, K. Helmerson and W. Phillips, 1996, in: Abstract booklet, 15th Int. Conf. on Atomic Physics.
Stephan, K.D., 1993, IEEE Antennas Propag. Mag. **35**, 13.
Stingl, A., Ch. Spielmann, F. Krausz and R. Szipöcs, 1994, Opt. Lett. **19**, 204.
Støvneng, J.A., and E.H. Hauge, 1989, J. Stat. Phys. **57**, 841.
Tanaka, M., 1989, Plasma Phys. and Controlled Fusion **31**, 1049.
Thomas, J.E., 1994, Contemp. Phys. **35**, 257.
van Tiggelen, B.A., A. Tip and A. Lagendijk, 1993, J. Phys. A **26**, 1731.
Wang, Y.P., and D.L. Zhang, 1995, Phys. Rev. A **52**, 2597.
Wigner, E.P., 1955, Phys. Rev. **98**, 145.
Yablonovitch, E., 1993, J. Opt. Soc. Am. B **10**, 283, and references therein, and other articles in this special issue on photonic bandgaps.
Zhu, S., A.W. Yu, D. Hawley and R. Roy, 1986, Am. J. Phys. **54**, 601.