

Causal wave propagation for relativistic massive particles: physical asymptotics in action

M V Berry

H H Wills Physics Laboratory, Tyndall Avenue, Bristol BS8 1TL, UK

E-mail: asymptotico@physics.bristol.ac.uk

Received 17 November 2011, in final form 5 December 2011

Published 10 January 2012

Online at stacks.iop.org/EJP/33/279

Abstract

Wavepackets representing relativistic quantum particles injected into a half-space, from a source that is switched on at a definite time, are represented by superpositions of plane waves that must include negative frequencies. Propagation is causal: it is a consequence of analyticity that at time t no part of the wave has travelled farther than ct , corresponding to the front of the signal. Nevertheless, interference fringes behind the front travel superluminally. For Klein–Gordon and Dirac wavepackets, the spatially integrated density increases because current is injected at the boundary. Even in the simplest causal model, understanding the shape of the wave after long times is an instructive exercise in the asymptotics of integrals, illustrating several techniques at a level suitable for graduate students; different spatial features involve contributions from a pole and from two saddle points, the uniform asymptotics for the pole close to a saddle, and the coalescence of two saddles into the Sommerfeld precursor immediately behind the front.

(Some figures may appear in colour only in the online journal)

1. Introduction

Long-standing interest in the motion of relativistic quantum particles has been rekindled by the discovery that neutrinos have mass, and the recent suggestion [1] that neutrinos might travel faster than the light speed c . An important feature is that even though waves can contain features that travel superluminally, these can never be used to send signals, so relativistic causality is not violated. This point was much discussed in the early 20th century, soon after the discovery of special relativity, in the context not of relativistic particles with mass but in optical materials where the phase and even group velocities can exceed c in frequency ranges close to an absorption band. The outcome of these researches [2–4] was that consistency with relativistic causality is a consequence of analyticity, and that understanding the detailed

evolution of wavepackets, especially for long times, is a tricky exercise in the asymptotics of integrals.

My purpose here is to show how the same ideas apply to relativistic particles with mass m , for which the energy E and momentum p are related by

$$E = \sqrt{(cp)^2 + (mc^2)^2}. \quad (1.1)$$

This is the analogue of a dispersive optical medium for which, in contrast to a dielectric material, there is no absorption. Questions of signalling and causality make no sense for packets with no definite beginning (e.g. the plane waves or Gaussians that have been studied elsewhere [5, 6]). Therefore in section 2 we follow [3, 4] and formulate quantum evolution as a boundary problem, in which particles, represented by a scalar wavefunction $\psi(x, t)$, are injected into the half-space $x > 0$, with source history prescribed on the boundary as a function $\psi(0, t)$ that is zero for $t < 0$. Such waves must have negative energies, raising interesting questions [7, 8], to which we return in the concluding section 7. Causality is demonstrated by reproducing the standard argument [8] based on analyticity in an integral representation for ψ .

Nowadays it is easy and instructive to compute the evolution of these packets numerically. The simplest representation involves the propagator [8], whose derivation is reproduced in section 3 and appendix A.

The physical significance of the scalar ψ is different for Klein–Gordon and Dirac particles. This is explained in section 4, and illustrated by the integrated density and the current for these two cases.

A selection of evolutions of ψ for different times and different boundary histories is displayed in section 5. The front—that is, the signal—moves with speed c , but certain features within the packet move superluminally.

Section 6 is a detailed analysis of the structure and evolution of these waves. For long times, the calculations require techniques from the asymptotic evaluation of integrals [9]: the method of stationary phase, including modifications for a saddle point in close proximity to a pole, and the breakdown of the method when the phase does not possess a narrow extremum. Different asymptotic techniques describe different regions of the evolving wave: near the boundary, in the main body of the wave and near the front. The analysis is an instructive illustration of the applicability of this area of mathematics—‘physical asymptotics’—at a level suitable for graduate students or advanced undergraduates.

Before getting into the advanced formalism, here is an elementary illustration, (suitable for beginning undergraduates or even high-school students), of the fact that features of waves can travel superluminally without violating causality. Consider two waves travelling at the speed of light (that is, representing massless particles), one forwards, one backwards, and with wavenumbers k and αk :

$$\psi_+ = \exp\{ik(x - ct)\}, \quad \psi_- = \exp\{i\alpha k(x + ct)\} \quad (0 < \alpha < 1). \quad (1.2)$$

Figure 1(a) shows the intensity pattern of their superposition, namely

$$I = |\psi_+ + \psi_-|^2 = 4 \cos^2 \left\{ \frac{1}{2} (1 - \alpha) k \left(x - \frac{1 + \alpha}{1 - \alpha} ct \right) \right\}. \quad (1.3)$$

The intensity fringes move with speed

$$v = \frac{1 + \alpha}{1 - \alpha} c, \quad (1.4)$$

which exceeds c , as can be also seen from the slope of the lines in the figure: their angle with the t axis exceeds 45° . A referee has pointed out that these superluminal fringes can be simulated as moiré fringes, made (on a computer screen or on transparent slides) by superposing two sets of equally spaced lines slightly rotated relative to each other.

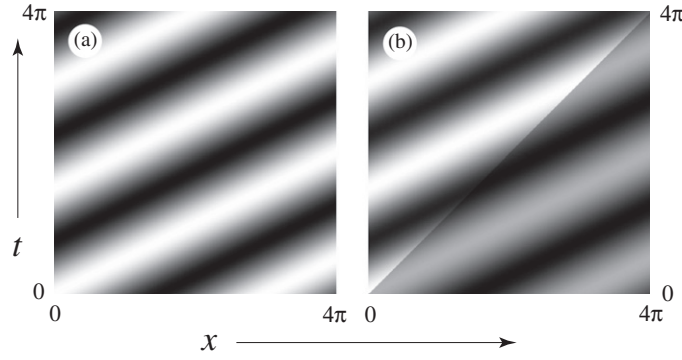


Figure 1. (a) Superposition of two waves (1.2) with $\alpha = 1/3$, each travelling with speed $c = 1$, showing maxima of intensity (1.3) travelling with speed $2c$. (b) With a modulation (signal) imposed on one of the waves ((1.5) with $\beta = 1/2$), the signal (discontinuity) travels with speed c .

Now impose a sudden modulation onto one of the waves, in an attempt to exploit this superluminal interference to send a signal:

$$\psi'_+ = \exp\{ik(x - ct)\}(1 + \beta\Theta(x - ct)), \quad (1.5)$$

where Θ denotes the unit step. As can be seen in figure 1(b), the attempt is unsuccessful: the modulation ignores the interference, and travels with speed c .

2. Causal wavefunction

It saves much writing to employ dimensionless variables. Convenient choices are to measure energy E in units of the rest energy mc^2 , momentum p in terms of the corresponding momentum mc , position X in terms of the reduced Compton wavelength, and time T in terms of the corresponding time:

$$E \equiv mc^2\omega, \quad p \equiv mck, \quad X \equiv \frac{\hbar}{mc}\xi, \quad T \equiv \frac{\hbar}{mc^2}t. \quad (2.1)$$

For electrons, the time unit \hbar/mc^2 is 1.3×10^{-21} s; for neutrinos with a sample rest energy of 1 eV, the unit is 6.6×10^{-16} s.

Thus the energy–momentum relation (1.1) becomes the dispersion relation

$$\omega = \sqrt{k^2 + 1}. \quad (2.2)$$

With this scaling, $\omega = 1$ corresponds to the rest energy, the speed of light is unity, and the phase and group velocities are

$$v_{\text{ph}} = \frac{\omega}{k} = \frac{\sqrt{k^2 + 1}}{k} = \frac{\omega}{\sqrt{\omega^2 - 1}}, \quad (2.3)$$

$$v_{\text{g}} = \frac{\partial\omega}{\partial k} = \frac{k}{\sqrt{k^2 + 1}} = \frac{1}{\omega\sqrt{\omega^2 - 1}} = \frac{1}{v_{\text{ph}}}.$$

As is well known, $v_{\text{ph}} > 1$ and $v_{\text{g}} < 1$ for real frequencies $1 \leq \omega < \infty$, and the nonrelativistic limit is $\omega \rightarrow 1$.

As we discuss further in section 4, the state vector may have one or more components (depending on the spin of the particles). For free particles, each component must satisfy the Klein–Gordon equation, and the scalar $\psi(x, t)$ that we consider here will represent one of them.

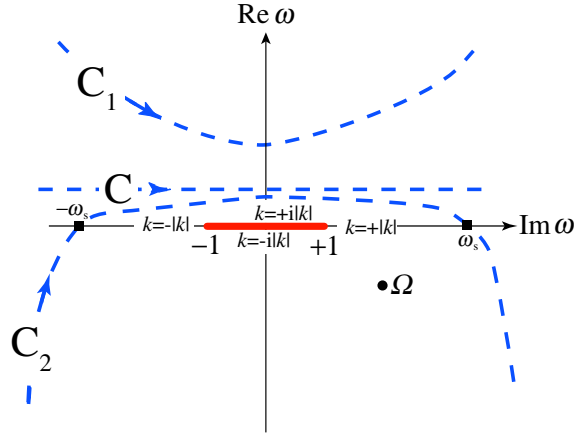


Figure 2. Complex plane of frequency (i.e. energy) ω , indicating branch cut (bold), phases of $k(\omega) = \sqrt{\omega^2 - 1}$ on the real axis, pole (bold circle), saddle points (bold squares), and integration contours C described in the text (dashed).

We will build $\psi(x, t)$ in the half-space $x > 0$ by superposing the corresponding plane-wave solutions, each with the phase

$$\frac{p}{\hbar}X - \frac{E}{\hbar}T = kx - \omega t = x\sqrt{\omega^2 - 1} - \omega t. \quad (2.4)$$

Thus

$$\psi(x, t) = \int_C d\omega A(\omega) \exp\{i(x\sqrt{\omega^2 - 1} - \omega t)\}, \quad (2.5)$$

in which the contour and the sign of the square root will be disambiguated soon. The Fourier amplitude $A(\omega)$ depends on the prescribed history of the wave at the boundary $x = 0$. Because we are interested in causality, we choose this as starting at $t = 0$, that is $\psi(0, t) = 0$ if $t < 0$. It follows that

$$A(\omega) = \frac{1}{2\pi} \int_0^\infty dt \psi(0, t) \exp(i\omega t). \quad (2.6)$$

This immediately implies that $A(\omega)$ is analytic in the upper half-plane $\text{Im } \omega > 0$. Moreover the contour C in (2.5) must run from $-\infty$ to $+\infty$ above the branch cut connecting the branch points $\omega = \pm 1$ (figure 2) of the square root momentum function $k(\omega) = \sqrt{\omega^2 - 1}$. The fact that the superposition unavoidably involves negative frequencies, and the interpretation of these as negative energies, raises problems of interpretation [8] that we defer to section 7. The function $k(\omega)$ is represented by a two-sheeted Riemann surface, and by choosing the physical sheet such that $k(\omega)$ is positive real for positive real ω (so that the packet represents waves travelling forwards) it follows that the phases elsewhere are as indicated in figure 2; and since with these phases both v_{ph} and v_{g} are positive for $\omega < 0$ as well as $\omega > 0$, the negative-frequency plane waves are also forward-propagating. Moreover, the momenta corresponding to $|\omega| < 1$ are positive imaginary, so these inevitably contributing nonclassical waves are evanescent for increasing x , as they must be.

To demonstrate causality, we note that

$$(x\sqrt{\omega^2 - 1} - \omega t) \rightarrow \omega(x - t) \text{ as } \omega \rightarrow \infty. \quad (2.7)$$

When $x > t$, the contour C can be deformed to infinity in the upper half-plane via C_1 in figure 2, and since it does not encounter the branch cut or any singularities of $A(\omega)$ the amplitude is zero:

$$\psi(x, t) = 0 \quad \text{if } 0 < x < t. \quad (2.8)$$

This standard argument [3, 8], based in analyticity, indeed establishes causality: for any signal entering the half-space at $t = 0$, no part of the packet reaches the position x before the time t —that is, the front of the signal cannot travel faster than light.

The above no-superluminal-signalling argument for massive relativistic particles in empty space is similar to the corresponding argument for light in a dielectric, in the sense that both involve analyticity. But in the latter case the prohibition of superluminal signalling also involves Maxwell's equations and—crucially—the Kramers–Kronig relations [10] expressing the requirement that the dielectric response is causal [11, 12], implying a connection between dispersion and absorption that is absent for propagation in vacuum.

3. Propagator

For any boundary history $\psi(0, t)$, it is not difficult to compute the wave for $x > 0$ numerically from the contour integral (2.5)—for example by deforming the contour to C_2 in figure 2 to optimize convergence, and using the NIntegrate routine in Mathematica. And as we will see in section 5 this representation in terms of frequencies ω is convenient for deriving approximations. For numerical evaluation, however, it is more efficient to use a time representation [7, 8], based on the propagator $P(x, t - t')$ in

$$\psi(x, t) = \int_0^\infty dt' \psi(0, t') P(x, t - t'). \quad (3.1)$$

P is the wave that evolves from the pulse $\psi(0, t) = \delta(t)$, corresponding to $A(\omega) = 1/2\pi$, so, from (2.5),

$$P(x, \tau) = \frac{1}{2\pi} \int_C d\omega \exp\{i(x\sqrt{\omega^2 - 1} - \omega\tau)\}. \quad (3.2)$$

Evaluation of this integral as explained in appendix A gives

$$P(x, \tau) = \delta(x - \tau) - \frac{x J_1(\sqrt{\tau^2 - x^2})}{\sqrt{\tau^2 - x^2}} \Theta(\tau - x), \quad (3.3)$$

in which J_1 is the standard Bessel function and Θ is the unit step.

In this way, the wave at any event (x, t) with $x < t$ can be calculated from

$$\psi(x, t) = \psi(0, t - x) - x \int_0^{t-x} dt' \psi(0, t') \frac{J_1(\sqrt{(t-t')^2 - x^2})}{\sqrt{(t-t')^2 - x^2}}. \quad (3.4)$$

This is the representation used in all numerical calculations to follow.

The wave propagating from $x = 0$ according to (3.4) is determined from the boundary history $\psi(0, t)$. A sufficiently general family of boundary histories, to be used in the illustrations that follow, is

$$\psi(0, t) = \Theta(t) t^N \exp(-i\Omega t) \quad \text{Re } \Omega > 0, \quad \text{Im } \Omega < 0. \quad (3.5)$$

This represents signals starting at $t = 0$ with an initial rise proportional to t^N , then oscillating with frequency $\text{Re } \Omega$ and eventually falling exponentially with decay exponent $\text{Im } \Omega$. (For $N = 0$ there is a mechanical implementation of this boundary history [13].) The corresponding Fourier amplitude is

$$A(\omega) = \frac{i^{N+1} N!}{2\pi (\omega - \Omega)^{N+1}}. \quad (3.6)$$

4. Current and density

The meaning of the scalar $\psi(x, t)$ depends on the physics of the wave it represents. The interpretations are different for Dirac and Klein–Gordon particles, but there is a common feature. In both cases, there is a density $\rho(x, t)$ and a current $j(x, t)$ (local expectation value of the velocity operator), related by a continuity equation

$$\partial_t \rho(x, t) = -\partial_x j(x, t). \quad (4.1)$$

However, for the class of causal waves we are considering the spatially integrated density

$$I(t) = \int_0^t dx \rho(x, t) \quad (4.2)$$

is not conserved: the waves are not normalized, because $I(t)$ increases as current is injected at the boundary $x = 0$. To express this precisely, we integrate (4.1) over x and t (noting that $\rho(x, t) = 0$ for $x > t$ because of causality), obtaining the relation

$$I(t) = \int_0^t dt' j(0, t'). \quad (4.3)$$

4.1. Klein–Gordon particles

With the scaling (2.1), this is

$$\partial_t^2 \psi - \partial_x^2 \psi + \psi = 0. \quad (4.4)$$

As is well known [7], and easy to verify directly, (4.1) is satisfied by

$$\rho(x, t) = -\text{Im} \psi^*(x, t) \partial_t \psi(x, t), \quad j(x, t) = \text{Im} \psi^*(x, t) \partial_x \psi(x, t). \quad (4.5)$$

($\rho(x, t)$ need not be positive-definite, but can be regarded as a type of charge density rather than a probability density; for an interesting perspective on the Klein–Gordon equation, see [14].) For the wave (3.5) with $N = 0$, the boundary current, computed from (3.4), is

$$j(0, t) = \exp(2t \text{Im} \Omega) \left(\text{Re} \Omega - \int_0^t d\tau \frac{J_1(\tau)}{\tau} \sin(\tau \text{Re} \Omega) \exp(-\tau \text{Im} \Omega) \right). \quad (4.6)$$

Figure 3(a) shows the evolution of the boundary current and integrated density for a typical case; the integrated density increases from $t = 0$, and saturates as the incident signal decays.

4.2. Dirac particles

In one dimension, the Dirac equation, involving a state vector with two components [6, 15] $\psi(x, t)$, $\phi(x, t)$ (because spin is irrelevant), each satisfying the Klein–Gordon equation (4.4), can be written as

$$\partial_t \begin{pmatrix} \psi(x, t) \\ \phi(x, t) \end{pmatrix} = (\sigma_x \hat{p} + \sigma_z) \begin{pmatrix} \psi(x, t) \\ \phi(x, t) \end{pmatrix} = \begin{pmatrix} 1 & -i\partial_x \\ -i\partial_x & -1 \end{pmatrix} \begin{pmatrix} \psi(x, t) \\ \phi(x, t) \end{pmatrix}. \quad (4.7)$$

Here σ_x and σ_z are the Pauli matrices, and \hat{p} is the momentum operator. The corresponding density and current, related by (4.1), are [7]

$$\begin{aligned} \rho(x, t) &= |\psi(x, t)|^2 + |\phi(x, t)|^2, \\ j(x, t) &= \begin{pmatrix} \psi^*(x, t) & \phi^*(x, t) \end{pmatrix} \sigma_x \begin{pmatrix} \psi(x, t) \\ \phi(x, t) \end{pmatrix} = 2 \text{Re} (\psi^*(x, t) \phi(x, t)). \end{aligned} \quad (4.8)$$

If the upper component $\psi(x, t)$ is specified, the lower component $\phi(x, t)$ can be calculated from the Dirac equation (4.7) by integrating over t , with the condition that $\phi(x, t \rightarrow \infty) = 0$,

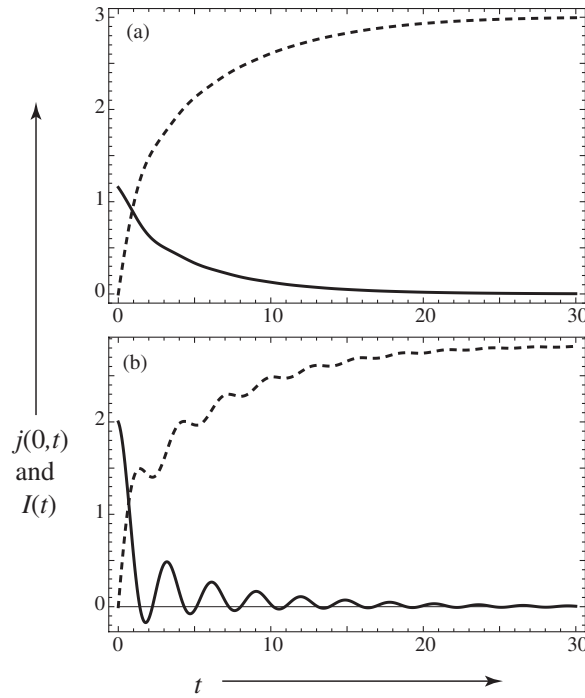


Figure 3. (a) Evolution of boundary current $j(0, t)$ (full curve) and integrated density $I(t)$ (dashed curve), for the Klein–Gordon wave propagating from the boundary history (3.5) with $N = 0$ and $\Omega = 2/\sqrt{3} - i/10$, calculated from (4.6); (b) as (a) for the Dirac wave, calculated from (4.9).

and then the current can be calculated from (4.8). For the wave with boundary history (3.5) with $N = 0$, this gives, after some calculation,

$$j(0, t) = 2 \operatorname{Re} \left[\frac{\Omega}{\Omega + 1} \exp(2t \operatorname{Im} \Omega) + \frac{\exp(it(\Omega^* + 1))}{\Omega + 1} + \frac{i}{\Omega + 1} \int_0^t d\tau \frac{J_1(\tau)}{\tau} (\exp(i\Omega\tau) \exp(2t \operatorname{Im} \Omega) - \exp(-i\tau) \exp(it(\Omega^* + 1))) \right]. \quad (4.9)$$

Figure 3(b) shows the evolution of the boundary current and integrated density for a typical case. Now the increase and saturation of the integrated density are accompanied by oscillations. These reflect the ‘zitterbewegung’ motion [6, 7], in which the expectation of the velocity operator σ_x in (4.8) oscillates between $\pm c$ because the eigenvalues of σ_x are ± 1 . This refers to the usual expectation, in which (4.8) is integrated over all x , not the local current considered here, for which the oscillations are damped.

5. Illustrations of relativistic scalar wavefunction

Figure 4 illustrates the spacetime structure of the wave for a typical case. Here and hereafter we display $\operatorname{Re} \psi(x, t)$ and $|\psi(x, t)|$; the behaviour of $\operatorname{Im} \psi(x, t)$ is similar, and in any case follows (up to sign) from that of $\operatorname{Re} \psi(x, t)$ and $|\psi(x, t)|$. For the example shown, $N = 0$, corresponding to a signal that is suddenly switched on with finite strength. Just as in the elementary example of figure 1(a), the interference fringes have slopes exceeding 45° relative

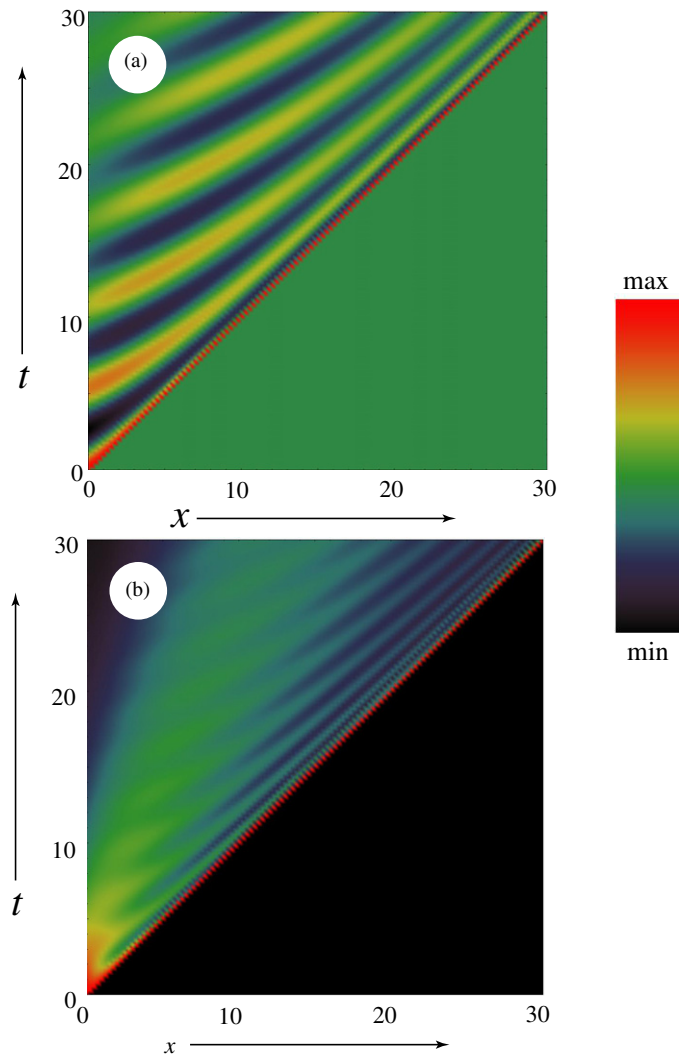


Figure 4. Spacetime density plots of (a) $\text{Re } \psi(x, t)$, (b) $|\psi(x, t)|$, for boundary history (3.5) with $N = 0$ and $\omega = 2/\sqrt{3} - i/10$.

to the t axis, so these features evolve superluminally, reflecting the fact the phase velocity $v_{\text{ph}} > c$ (cf (2.3)). We emphasize that this does not imply a violation of relativistic causality, but rather a reshaping of the progressing waveform, in which the features try to catch up with the (causally propagating) front but fade out before reaching it. An analogy is with gravity waves on deep water, for which $v_{\text{ph}} = 2v_g$, explaining why people in a rowing-boat, seeing the approach of waves generated by a big ship, have twice as long as they might think before the disturbance reaches them.

Figure 5 shows the waves as functions of x/t for three different fixed times t and two different complex carrier frequencies Ω . In both cases $\text{Re } \Omega = 2/\sqrt{3}$, which from (2.3) would correspond to $v_g(\text{Re } \Omega) = 1/2$. On the left panels (figures 5(a), (c) and (e)), $N = 0$, so the sudden signal front at $x/t = 1$ is a finite jump, and $\text{Im } \Omega = 0.01$ is small, so the signal tail decays slowly towards $x = 0$. On the right panels (figures 5(b), (d) and (f)), $N = 3$, so the signal

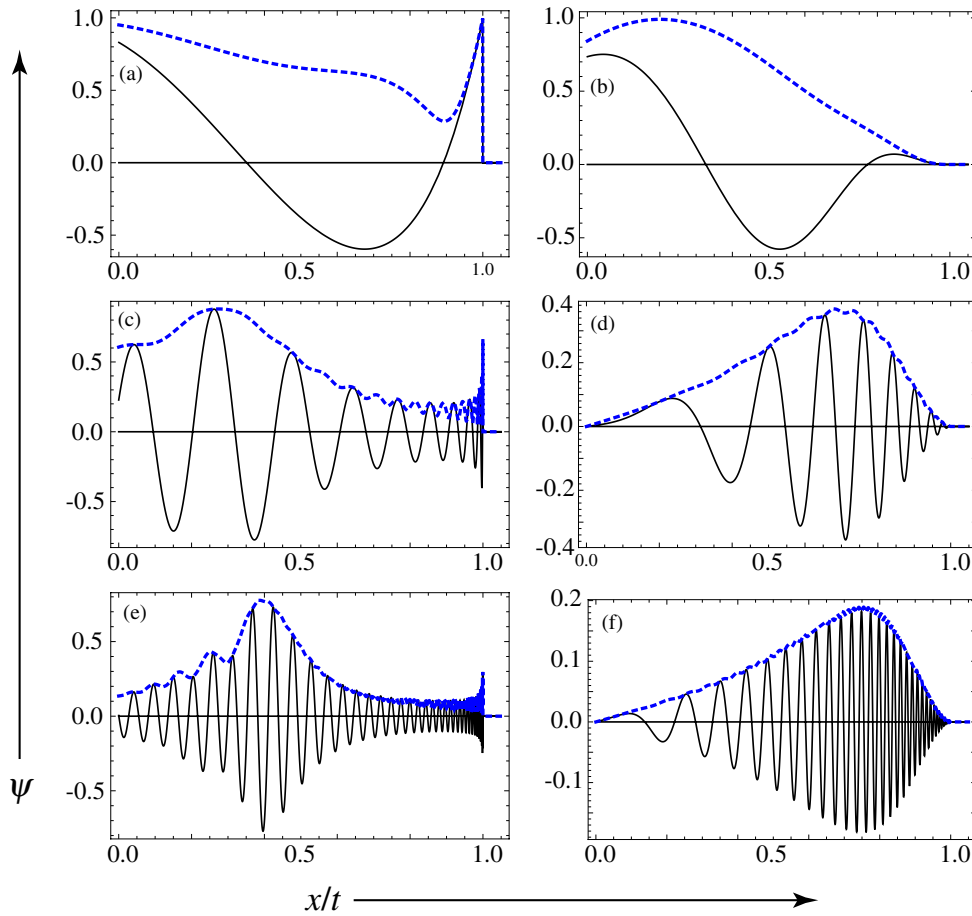


Figure 5. Plots of $\text{Re } \psi(x, t)$ (full curves) and $|\psi(x, t)|$ (dashed curves) for boundary condition (4.1) with (a), (c), (e) $N = 0$, $\Omega = 2/\sqrt{3} - i/100$; (b), (d), (f) $N = 3$, $\Omega = 2/\sqrt{3} - i$, and (a), (b) $t = 5$; (c), (d) $t = 50$; (e), (f) $t = 200$.

rises smoothly behind the front at $x/t = 1$, and $\text{Im } \Omega = 1$ is much larger, so that the signal tail decays more rapidly towards $x = 0$. For both values of Ω , the maximum signal amplitude for long times occurs for $x/t < \text{Re } v_g(\Omega)$ ($= 0.5003$ for the right panels and 0.992 for the left panels)—that is for smaller x/t than would be predicted by a naive group-velocity argument. The frequency of the local oscillations in these wavepackets is greatest near the front, which is dominated by the large plane-wave frequencies ω that travel at nearly light speed, and slowest near $x = 0$, which corresponds to plane waves representing particles moving more slowly. To understand these features in quantitative detail requires asymptotics, which will now be explained.

6. Long-time asymptotics

The form of the wavepacket, which is non-zero between $x = 0$ and $x = t$, changes with time. From the scaling (2.1), and the values of the time unit immediately following (2.1), any

application to real particles will involve $t \gg 1$. Therefore we should understand the long-time asymptotics of the wavepacket. For this purpose, it is convenient to replace x by the variable

$$u \equiv \frac{x}{t} \quad (6.1)$$

which has the range $0 \leq u \leq 1$ and can be regarded as a velocity. And it will suffice to consider the boundary histories (3.5) with $N = 0$. Then the evolving wave is

$$\psi(ut, t) = \frac{i}{2\pi} \int_C \frac{d\omega}{\omega - \Omega} \exp\{it(u\sqrt{\omega^2 - 1} - \omega)\}. \quad (6.2)$$

In the asymptotics to follow, we will consider t as a large parameter, and obtain a series of approximations. These will be illustrated by the successive frames in figure 6 for $t = 50$ and $\Omega = 2/\sqrt{2} - i/100$ —values chosen to display the different phenomena clearly. Figure 6(a) shows the exact (to graphical accuracy) wave for this case, numerically computed from (3.4).

6.1. Contribution from pole

If the right-hand part of the contour C_2 in figure 2 is shifted to the left, it will pass the pole at $\omega = \Omega$, whose contribution is the residue in (6.2), counted negatively:

$$\psi_{\text{pole}}(ut, t) = \exp\{it(u\sqrt{\Omega^2 - 1} - \Omega)\}. \quad (6.3)$$

This is just the input signal (3.5) propagating as a complex-frequency plane wave, with dispersion neglected. The propagation speed of this part of the wave is the phase velocity (2.3) corresponding to Ω , which exceeds unity; this is the superluminal propagation visible as the sloping fringes in figure 4. For massless particles, that is with $\sqrt{\omega^2 - 1}$ replaced by ω , that is, where there is no dispersion, ψ_{pole} would be the exact solution. Note that the contribution from the pole is $O(t^0)$.

6.2. Contributions from saddles

Along the real axis, the phase in (6.2) oscillates rapidly except at the saddle points $\omega = \pm\omega_s$ where the phase is stationary:

$$\omega_s = \frac{1}{\sqrt{1 - u^2}}. \quad (6.4)$$

These frequencies have the physical meaning that they correspond to local plane waves with the group velocity needed to reach the event (x, t) from $(0, 0)$. Then the standard stationary-phase method [9, 16] (quadratic approximation of the phase around $\omega = \pm\omega_s$) can be applied as outlined in appendix B, giving the contribution from these points as

$$\psi_{\text{saddle}\pm}(ut, t) = \frac{u \exp\{\pm i(\frac{1}{4}\pi - t\sqrt{1 - u^2})\}}{\sqrt{2\pi t}(1 - u^2)^{1/4}(1 \mp \Omega\sqrt{1 - u^2})}. \quad (6.5)$$

Since $\text{Re } \Omega > 0$, the denominator implies that the contribution from $+\omega_s$ is larger than that from $-\omega_s$, so we call $+\omega_s$ the dominant saddle. Note that the contribution from the saddle is $O(1/\sqrt{t})$.

Figure 6(b) is a comparison of the exact ψ with the pole and dominant saddle contributions. It is clear that the pole contribution is a good approximation for small x/t . This corresponds to late times, when all transients have disappeared and the signal injected into the half-space has become established at positions away from the front. By contrast, the dominant saddle is a good approximation for larger values of x/t , where the injected signal has not been established and the local oscillations reflect the local group velocity. At the crossover between these approximations, in this example near $u = 1/2$, corresponding to the group velocity of the injected frequency $\text{Re } \Omega$, the saddle formula possesses a large unphysical spike, indicating that in this region the pole and saddle contributions cannot be separated.

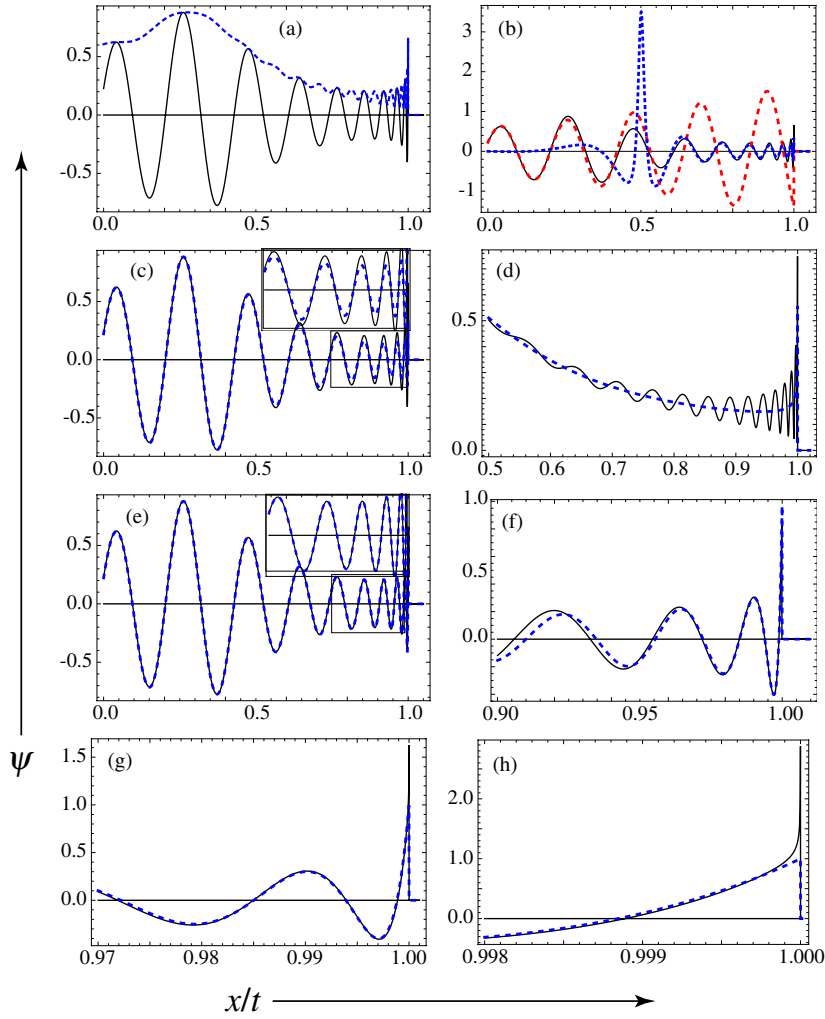


Figure 6. Plots of wavefunction $\psi(x, t)$ for boundary condition (3.5) with $N = 0$ and $\Omega = 2/\sqrt{3} - i/100$, as functions of $u = x/t$ for $t = 50$. (a) $\text{Re } \psi$ (full curve) and $|\psi|$ (dashed curve), computed exactly from (3.4) (this repeats figure 5(c)); (b) exact $\text{Re } \psi$ (full curve), pole contribution $\text{Re } \psi_{\text{pole}}$ from (6.3) (dashed), and dominant saddle contribution $\text{Re } \psi_{\text{saddle+}}$ from (6.5) (dotted); (c) exact $\text{Re } \psi$ (full curve) and uniform approximation $\text{Re } \psi_{\text{pole,saddle+}}$ from (6.13) (dashed); (d) exact $|\psi|$ (full) and uniform approximation $|\psi_{\text{pole,saddle+}}|$ (dashed), over the range $0.5 \leq x/t \leq 1$, showing deviations; (e) exact $\text{Re } \psi$ (full curve) and uniform approximation plus the saddle w_- , i.e. $\text{Re}(\psi_{\text{pole,saddle+}} + \psi_{\text{saddle-}})$ (dashed); (f) exact $\text{Re } \psi$ (full curve) and Sommerfeld precursor $\text{Re } \psi_{\text{precursor}}$ from (6.17) (dashed); (g) $\text{Re}(\psi_{\text{pole,saddle+}} + \psi_{\text{saddle-}})$ (full) and Sommerfeld precursor $\text{Re } \psi_{\text{precursor}}$ (dashed); (h) magnification near the front, showing $\text{Re}(\psi_{\text{saddle+}} + \psi_{\text{saddle-}})$ (full) and Sommerfeld precursor $\text{Re } \psi_{\text{precursor}}$ (dashed).

6.3. Uniform approximation for pole and saddle combined

To get a precise asymptotic description of the wave, incorporating the contributions from the pole and the dominant saddle in a consistent manner, we use the technique of uniform approximation [9, 17]. In this method, the integral to be evaluated is mapped onto the simplest canonical form: a comparison integral with the same topology of relevant saddles

and singularities. The best-known example is the class of integrals with two saddles that can coalesce as a parameter is varied [18], for which the comparison integral is the Airy function. For our integral (6.2), we have a saddle that can come close to a pole as a parameter (in this case u) varies. The obvious comparison integral, involving the complementary error function, is (cf equations (7.7.2) and (7.2.3) of [17])

$$\int_{-\infty}^{\infty} \frac{dW}{W - W_{\Omega}} \exp\left(-\frac{1}{2}itW^2\right) = -i\pi \exp\left(-\frac{1}{2}itW_{\Omega}^2\right) \operatorname{erfc}\left\{-\exp\left(-\frac{1}{4}i\pi\right)\sqrt{\frac{t}{2}}W_{\Omega}\right\}, \quad (6.6)$$

in which the saddle is at $\Omega = 0$ and the parameter is the position W_{Ω} of the pole in the W plane.

To represent (6.2) in terms of (6.6), the mapping between the integration variables ω and W must be one-to-one. Therefore the saddles must coincide, i.e.

$$\begin{aligned} u\sqrt{\omega^2 - 1} - \omega &\equiv -\sqrt{1 - u^2} - \frac{1}{2}W^2, \\ \omega = \omega_s &= \frac{1}{\sqrt{1 - u^2}} \Leftrightarrow W = 0. \end{aligned} \quad (6.7)$$

It follows that the location of the pole in the W plane is

$$W_{\Omega} = \sqrt{2(\Omega - u\sqrt{\Omega^2 - 1} - \sqrt{1 - u^2})}. \quad (6.8)$$

Thus the wave becomes (after neglecting the contribution from the subdominant saddle $\omega = -\omega_s$, which is far from the pole and so plays no part in the uniform approximation)

$$\psi(ut, t) \approx \frac{i}{2\pi} \exp(-t\sqrt{1 - u^2}) \int_{-\infty}^{\infty} dW g(W) \exp\left(-\frac{1}{2}tW^2\right). \quad (6.9)$$

We write the multiplier function $g(W)$, incorporating the pole and the Jacobian of the transformation of variables, in the following convenient form, in which the pole and saddle are separated:

$$g(W) = \left(\frac{\partial W}{\partial \omega}\right)^{-1} \frac{1}{\omega - \Omega} = \frac{A}{W - W_{\Omega}} + B + W(W - W_{\Omega})h(W). \quad (6.10)$$

Thus

$$\begin{aligned} A &= \lim_{\omega \rightarrow \Omega} \left(\frac{\partial W}{\partial \omega}\right)^{-1} \frac{W - W_{\Omega}}{\omega - \Omega} = 1, \\ B &= \frac{1}{W_{\Omega}} + \lim_{\omega \rightarrow \omega_s} \left(\frac{\partial W}{\partial \omega}\right)^{-1} \frac{1}{\omega - \Omega} = \frac{1}{W_{\Omega}} + \frac{u}{(1 - u^2)^{3/4}(\omega_s - \Omega)}, \end{aligned} \quad (6.11)$$

in which the second equality for B is obtained by differentiating the map (6.7) twice, leading to

$$\left(\frac{\partial W}{\partial \omega}\right)_{\omega \rightarrow \omega_s}^2 = \frac{u}{(\omega_s^2 - 1)^{3/2}} = \frac{(1 - u^2)^{3/2}}{u^2}. \quad (6.12)$$

The final step is to note that the term involving $h(W)$ vanishes at the saddle and pole which dominate the integral. This term would contribute to higher-order asymptotics, so we neglect it in this simplest treatment. Of the integrals that remain, the one involving the coefficient A

is of the form (6.6), and that involving B is elementary. Thus we find the uniform pole–saddle approximation

$$\begin{aligned} \psi_{\text{pole,saddle+}}(ut, t) = \exp(-it\sqrt{1-u^2}) \left[\frac{\exp(\frac{1}{4}i\pi)}{\sqrt{2\pi t}} \left(\frac{1}{W_\Omega} + \frac{u}{(1-u^2)^{3/4}(1-\Theta\sqrt{1-u^2})} \right) \right. \\ \left. + \frac{\exp(-\frac{1}{2}itW_\Omega^2)}{2} \operatorname{erfc} \left\{ -\exp\left(-\frac{1}{4}i\pi\right) \sqrt{\frac{t}{2}W_\Omega} \right\} \right]. \end{aligned} \quad (6.13)$$

As figure 6(c) shows, this captures the pole-to-saddle transition near $u = 0.5$ very accurately. Nevertheless, there are some small oscillatory deviations, which get stronger closer to the front; they can be seen more clearly in the inset to figure 6(c) and in figure 6(d), which is a magnified graph of the modulus $|\psi|$. The deviations result from neglect of the contribution from the subdominant saddle $\omega = -\omega_s$; when this is included (figure 6(e)), the agreement improves substantially (see also the inset to compare with figure 6(c)).

6.4. Near the front: the Sommerfeld precursor

Very close to the front, that is $1-u \ll 1$ in (6.2), the approximations based on saddles must fail, because $\omega_s \gg 1$ near the front (cf (6.4)), and the phase no longer varies quadratically. Instead, the relevant ω dependence is

$$\sqrt{\omega^2 - 1} \approx \omega - \frac{1}{2\omega}. \quad (6.14)$$

Thus the approximate form of ψ near the front—the ‘Sommerfeld precursor’ [3]—requires evaluation of

$$\psi_{\text{precursor}}(ut, t) = \frac{i}{2\pi} \int_C \frac{d\omega}{\omega} \exp \left\{ -it \left(\omega(1-u) + \frac{1}{2\omega} \right) \right\}. \quad (6.15)$$

Following Sommerfeld, we deform the contour C_2 in figure 2 to a large circle of radius R surrounding the pole, parameterized by an angle θ . The corresponding transformation of variables,

$$u = R \exp(i\theta) \quad R = \sqrt{\frac{1}{2(1-u)}}, \quad 0 \leq \theta < 2\pi, \quad (6.16)$$

reduces the integral (6.15) to a standard representation of a Bessel function. Thus the precursor is

$$\begin{aligned} \psi_{\text{precursor}}(ut, t) &= \frac{1}{2\pi} \int_0^{2\pi} d\theta \exp\{-it\sqrt{2(1-u)} \cos \theta\} \\ &= J_0(t\sqrt{2(1-u)}). \end{aligned} \quad (6.17)$$

Note that this contribution is $O(t^0)$; for the more general boundary history (3.5), it would be $O(t^N)$.

As the magnification in figure 6(f) shows, this is a very accurate representation of the wave near the front. By contrast, the uniform pole–saddle approximation, plus the contribution from the subdominant saddle, fails at the front because it diverges there; see the further magnification in figure 6(g). In fact the uniform approximation is irrelevant close to the front; the sum of the two saddle contributions, each of which diverges at the front (cf (6.5)), matches smoothly onto the Bessel function in the precursor, because of the asymptotic relation

$$\begin{aligned} J_0(t\sqrt{2(1-u)}) &\approx \frac{\exp\{i(\frac{1}{4}\pi - t\sqrt{2(1-u)})\}}{\sqrt{2\pi t}(2(1-u))^{1/4}} + \frac{\exp\{-i(\frac{1}{4}\pi - t\sqrt{2(1-u)})\}}{\sqrt{2\pi t}(2(1-u))^{1/4}} \\ &= \psi_{\text{saddle+}}(ut, t) + \psi_{\text{saddle-}}(ut, t) \text{ for } 1-u \ll 1 \text{ and } t\sqrt{1-u} \gg 1. \end{aligned} \quad (6.18)$$

This connection, and the divergence of the contribution from the two saddles, is illustrated in the final magnification in figure 6(h).

7. Concluding remarks

It is clear that in the asymptotic regime $t \gg 1$ a separation of phenomena emerges: different features of the wave are associated with different regions in the complex plane of the integration variable ω . For small values of $u = x/t$, including the region near the boundary $x = 0$, the pole at $\omega = \Omega$ dominates (section 6.1), and phase features of the wave travel superluminally. For larger u , roughly, $u > \text{Re} \sqrt{1 - \Omega^{-2}}$, the wave is well approximated (section 6.2) by the dominant saddle (6.4). The smooth matching between these regimes is achieved by pole-to-saddle uniform asymptotics (section 6.3), but for further accuracy the contribution from the subdominant saddle, incorporating the negative frequencies in the wave, must be included. All these approximations fail near the front, where a different contour deformation leads to the Sommerfeld precursor (section 6.4).

Consistent with the separation of phenomena associated with different regions of the integrand in (6.2), we might expect contributions from the branch points at $\omega = \pm 1$. However, a scaling argument indicates that these contributions would be $O(1/t^2)$ and therefore small in comparison with $O(t^0)$ from the pole and $O(1/\sqrt{t})$ from the saddles.

Now we return to the negative frequencies in the superposition (2.5). The fact that these are unavoidable in relativistic causal wavepackets brings into sharp focus the much-discussed question [7, 8] of the meaning of negative energies. To investigate how important these are, we note that the negative frequencies between $\omega = 0$ and $\omega = -1$ correspond to waves that are evanescent (that is, decaying for increasing x), so the more problematic frequencies are those in the range $\omega < -1$, that is, energies $E < -mc^2$. The fraction of the total signal power in these negative-frequency waves is

$$r = \frac{\int_{-\infty}^{-1} d\omega |A(\omega)|^2}{\int_{-\infty}^{+\infty} d\omega |A(\omega)|^2}. \quad (7.1)$$

We estimate this fraction for a simpler boundary-history model than (3.5), namely a signal with real carrier frequency Ω that is suddenly switched on and off at scaled times $t = 0$, $t = t_f$. A short calculation leads to

$$r = \frac{2}{\pi t_f} \int_1^{\infty} d\omega \frac{\sin^2((\omega + \Omega)t_f)}{(\omega + \Omega)^2} \approx \frac{1}{\pi t_f(\omega + 1)} = \frac{h}{\pi T_f(E_k + 2mc^2)} \leq \frac{h}{2\pi T_f mc^2}. \quad (7.2)$$

Here T_f is the unscaled duration of the signal and E_k is the kinetic energy of the particles (i.e. the total energy is $E_k + mc^2$). This is small for any realistic signal, in which the bandwidth is small—or, in particle language, in which the duration of the signal is large compared with the Compton time h/mc^2 .

For electrons, there seems to be a consensus that there is no consistent interpretation of the negative-energy solutions within a one-particle quantum theory. But one often-discussed possibility [7] is that the frequencies $\omega < -1$ correspond to positrons with energies $E > +mc^2$. At first sight, this interpretation seems far-fetched for the causal waves we are discussing here. It would imply that every time an electron microscope is switched on a fraction r of the electrons would appear as positrons, which promptly annihilate with some of the positive-energy electrons, generating a burst of γ -rays. Nevertheless, it is instructive to estimate the total number N of such positrons that would be generated during the signal duration T_f . This is r times the electron number current times T_f , which from (7.2) gives the interesting result that N is approximately the number of electrons entering the microscope during a Compton time.

A typical electron microscope electric current is 1 nA, corresponding to about 10^9 electrons per second. This gives, with (7.2) and noting that T_{τ} cancels, $N \sim 10^{-13}$ positrons—interesting in principle, but not likely to influence the working practices of electron microscopists.

Appendix A. Derivation of propagator (3.4)

By separating the large ω behaviour, we can write (3.3) in the form

$$P(x, \tau) = \delta(x - \tau) - Q(x, \tau) \Theta(\tau - x), \quad (\text{A.1})$$

in which

$$Q(x, \tau) = \frac{1}{2\pi} \int_C d\omega \exp(-i\omega\tau) (\exp\{i(x\sqrt{\omega^2 - 1})\} - \exp(ix)). \quad (\text{A.2})$$

The contour C can be deformed into C_2 in figure 2, and then further deformed to surround the branch cut, because the integral contains no poles. The integral involving the factor $\exp(ix)$ does not depend on the multivalued function $r(\omega)$, so its contribution vanishes. Thus

$$\begin{aligned} Q(x, \tau) &= -\frac{1}{\pi} \int_{-1}^1 d\omega \exp(-i\omega\tau) \sinh(x\sqrt{1 - \omega^2}) \\ &= -\frac{2}{\pi} \int_0^{\pi/2} d\theta \cos\theta \cos(\tau \sin\theta) \sinh(x \cos\theta), \end{aligned} \quad (\text{A.3})$$

the second equality follows from the transformation $\omega = \sin\theta$. The θ integral can be evaluated using an elementary modification of the formula (10.9.13) in [17], and (3.4) follows.

Appendix B. Derivation of (6.5)

The standard stationary-phase method [9] is an approximation for

$$I = \int_a^b d\omega g(\omega) \exp(itf(\omega)) \quad (\text{B.1})$$

in which t is large and $f(\omega)$ has a stationary point in the integration range: $\partial_{\omega}f(\omega_s) = 0$ and $a < \omega_s < b$. The method is based on the observation that the main contribution to I comes from the neighbourhood of ω_s because when t is large the integrand is rapidly oscillating elsewhere and the contributions away from ω_s cancel. Expanding the exponent to second order about ω_s and extending the integration range to the whole real axis (i.e. ignoring the contributions from the end-points a and b which are smaller), and approximating $g(\omega)$ by its value $g(\omega_s)$ at the stationary point gives a Gaussian integral and the standard [9] result

$$I = \frac{g(\omega_s)}{\sqrt{it\partial_{\omega}^2 f(\omega_s)}} \exp\left(i\left(tf(\omega_s) + \frac{1}{4}\pi \operatorname{sgn}\partial_{\omega}^2 f(\omega_s)\right)\right). \quad (\text{B.2})$$

When applied to the integral (6.2) with the saddle (6.4), this leads straightforwardly to the result (6.5) and this exercise is a useful illustrative application of the stationary-phase technique.

References

- [1] Adam T *et al* 2011 Measurement of the neutrino velocity with the OPERA detector in the CNGS beam arXiv:1109.4897
- [2] Brillouin L and Sommerfeld A 1960 *Wave Propagation and Group Velocity* (New York: Academic)
- [3] Sommerfeld A 1950 *Optics: Lectures on Theoretical Physics* vol 4 (New York: Academic)
- [4] Jackson J D 1975 *Classical Electrodynamics* (New York: Wiley)
- [5] Demikhovskii V Y, Maksimova G M, Perov A A and Frolova E V 2010 Space-time evolution of Dirac wave packets *Phys. Rev. A* **82** 052115

- [6] Thaller B 2004 Visualising the kinematics of Dirac wave packets <http://www.uni-graz.at/imawww/vqm/articles.html>
- [7] Messiah A 1962 *Quantum Mechanics* (Amsterdam: North-Holland)
- [8] Thaller B 1992 *The Dirac Equation* (Berlin: Springer)
- [9] Wong R 1989 *Asymptotic Approximations to Integrals* (New York: Academic)
- [10] Bohren C F 2010 What did Kramers and Kronig do and how did they do it? *Eur. J. Phys.* **31** 573–7
- [11] Toll J S 1956 Causality and the dispersion relation: logical foundations *Phys. Rev.* **104** 1760–70
- [12] Kinsler P 2011 How to be causal: time, spacetime and spectra *Eur. J. Phys.* **32** 1687–700
- [13] Mouchet A 2008 Interaction with a field: a simple integrable model with backreaction *Eur. J. Phys.* **29** 1033–50
- [14] Bertozzi E 2010 Hunting the ghosts of a ‘strictly quantum field’: the Klein–Gordon equation *Eur. J. Phys.* **31** 1499–515
- [15] Nitta H, Kudo T and Minowa H 1999 Motion of a wave packet in the Klein paradox *Am. J. Phys.* **67** 966–71
- [16] Dingle R B 1973 *Asymptotic Expansions: Their Derivation and Interpretation* (New York: Academic)
- [17] DLMF 2010 *NIST Handbook of Mathematical Functions* (Cambridge: Cambridge University Press) <http://dlmf.nist.gov>
- [18] Chester C, Friedman B and Ursell F 1957 An extension of the method of steepest descents *Math. Proc. Camb. Phil. Soc.* **53** 599–611