Extensions of Quantum Physics

Edited by Andrzej Horzela and Edward Kapuscik

Extensions of Quantum Physics

Edited by Andrzej Horzela and Edward Kapuscik



Published by C. Roy Keys Inc. 4405, rue St-Dominique Montreal, Quebec H2W 2B2 Canada http://redshift.vif.com

© C. Roy Keys Inc. 2002

First Published 2002

National Library of Canada Cataloguing in Publication Data

Main entry under title:

Extensions of quantum physics / edited by Andrzej Horzela and Edward Kapuscik.

Includes bibliographical references. ISBN 0-9683689-8-0

Quantum theory. 2. Tunneling (Physics).
 Horzela, Andrzej, 1952- II. Kapuscik, Edward, 1938-

QC174.12.E98 2002 530.12 C2002-901387-9

Cover design by Amanda Lawrence

Foreword

In 1999 we began holding the Cracow - Clausthal annual workshops devoted to unsolved fundamental problems of quantum mechanics. The workshops were initiated by Professor Heinz-Dietrich Doebner from the Arnold Sommerfeld Institute of Mathematical Physics and the Technical University of Clausthal at Clausthal-Zellerfeld (Germany), and by a our group at the Henryk Niewodniczański Institute of Nuclear Physics at Kraków (Poland). The first workshop was organized under the title *Tunneling Effect and Other Fundamental Problems of Quantum Physics* and was held in Kraków from November 22 to 28 of that year. During all the discussions the participants came to the conclusion that the framework of the workshops should be enlarged, and that the appropriate name for all future workshops would be *Extensions of Quantum Theory*. The second workshop was held in Kraków from October 12 to 15, 2000. The third, which was held from July 18 to 21, 2001 was associated with the 2nd International Symposium *Quantum Theory and Symmetries*, organized by us and hosted by the H. Niewodniczański Institute of Nuclear Physics.

Simultaneously, it was agreed that the results of the workshops should be published in a collection of regular articles in special volumes. It is our pleasure to present the first such a volume, which covers some topics discussed up to now. We are grateful to the publisher of Apeiron, C. Roy Keys, for providing us this opportunity.

The present volume starts with a discussion of superluminal signal velocities in tunneling experiments with microwaves, and a controversy connected with the locality problem. Then we continue with the problem of localization for photons, which also belongs to the list of unsolved problems of quantum theory. Closely related to these topics are the problems of time of arrival in quantum physics and preferred reference systems in Maxwell electrodynamics. A new look at the problem of superluminal velocities is presented on the basis of spacetimes with multidimensional times. In addition, the complicated problem of tunneling through many succesive bariers and many-layer systems is discussed. The utility of lesser-known representations of quantum physics in describing the tunneling effect is also presented.

Finally, we have decided to include in this volume a paper which discusses possible use of quantum non-commutative geometries for constructing more realistic quantum models of our Universe. In our view, newer emerging models will naturally incorporate more observational events and be fundamentally different from currently used models of spacetime. In particular, we believe that in such models all the fundamental cosmological observations which contradict the standard point of view will find their natural explanation.

Apart from fundamental problems, we also welcome new applications of standard theories. This is why we have included here some papers containing interesting results

ii

obtained in traditional frameworks.

We wish to express our deep gratitude to the Polish State Committee for Scientific Research for providing the funding that made it possible to hold all the workshops, and to all those who helped us in preparing the workshops, in particular to Professor Andrzej Budzanowski, Director General of the H. Niewodniczański Institute of Nuclear Physics and to our colleagues in the Institute. Last but not least, we extend special thanks to all the contributors to this volume. We also hope our readers will find this volume to be of interest, and that they will look forward to further installments.

December 2001,

Andrzej Horzela and Edward Kapuścik H. Niewodniczański Institute of Nuclear Physics

Kraków, Poland

Table of Contents

G. Nimtz, A. Haibel, A. A. Stahlhofen, and RM. Vetter On Universal Properties of Tunneling	1
G. C. Hegerfeldt	
Particle Localization and the Notion of Einstein Causality	11
A. S. Shumovsky Locality of Quantum Electromagnetic Radiation	19
J. Leon Time of Arrival in Classical and Quantum Mechanics	29
J. Rembieliński Superluminal Phenomena and the Phenomenological Maxwell Equations	45
V. S. Barashenkov and M. Z. Yuriev Sub- and Superluminal Velocities in Space with Vector Time	53
V. S. Olkhovsky, E. Recami, and G. Salesi Superluminal Tunneling through Two Successive Barriers	63
A. M. Gabovich, V. M. Rozenbaum, and A. I. Voitenko Tunneling in Two- and Three-layer Systems with Allowance for Dynamic Image Forces	77
B. Segev Tunneling in the Wigner Representation	93
M. Heller Geometric Structure of the Big Bang	103
L. Kostro	
The Quantities c^4/G and c^5/G and the Basic Equations of Quantum Mechanics	115
P. Machnikowski, P. Magnuszewski, and A. Radosz Stability of the Bell-shaped Excitations in Discrete Models	
of Molecular Chains	121

On Universal Properties of Tunneling

G. Nimtz, A. Haibel, A. A. Stahlhofen, and R.–M. Vetter II. Physikalisches Institut, Universität zu Köln, D-50937 Köln

Photonic tunneling violates Einstein causality. Superluminal signal and energy velocities have been measured in studies of photonic tunneling. The signal energy is always finite; thus as a consequence of quantum mechanics, signals must be frequency band limited. This result represents a fundamental physical property. We conjecture that tunneling time is universal.

Keywords: tunneling time, superluminal, signals

1. Introduction

First we will present some experimental results on photonic tunneling observed in microwave and optical experiments. Experiments have revealed superluminal signal and energy velocities [1, 2, 3]. According to the textbooks, Einstein or "strict" causality means neither signals nor energy can travel faster than c, the speed of light in vacuum. To resolve this sophisticated dilemma, the main part of the paper is devoted to a discussion of the properties of a signal and signal velocity. Surprisingly, it follows that Einstein causality may be violated by photonic tunneling. The effect can indeed follow the cause at superluminal velocity. However, this result does not include the possibility of changing the past. Constructing a time machine is yet not possible.

1.1 Photonic Tunnel Barriers

Three examples of photonic barriers are represented in Fig.1. showing an undersized waveguide between two normal guides, a periodic dielectric heterostructure (often called a one-dimensional photonic lattice), and a double prism with a gap of rarer refractive index acting as a photonic barrier. The latter set-up is described as frustrated total internal reflection (FTIR). Dispersion relations for the transmission of the lattice and the undersized waveguide are shown in the same figure.

Photonic barriers and wave mechanical barriers are characterized by a field mode solution with an imaginary wave number called the evanescent mode in classical optics. The evanescent field crossing the barrier decays exponentially with distance, however, *without* changing its phase. The IEEE, for instance, summarized this property in the following definition: *An evanescent mode in an undersized waveguide is a field configuration in a waveguide such that the amplitude of the field diminishes along the waveguide but the phase is unchanged.*



Fig. 1. Examples (a) of a waveguide with an undersized central part, (b) a onedimensional periodic dielectric hetero-structure, and (c) a double prism (FTIR) with an evanescent gap. The graphs below show the dispersion relations for transmission from structures (a) and (b); the double prism structure dispersion is qualitatively the inverse of example (a). The dispersion of the periodic heterostructure displays a forbidden gap which corresponds to a tunneling regime.

1.2 Observed Superluminal Signal and Energy Velocity

A single digital pulse is shown in Fig. 2. This signal has crossed a photonic barrier at the speed of 4.7c arriving at the observer 500 ps earlier than a waveguided copy thereof which travelled the same distance at the vacuum speed of light [4]. The observer received the tunnelled signal earlier, which means that the cause to effect gap has been shortened.

Note, that the tunnelled signal is not markedly reshaped. This is due to its frequency band limitation and the fact that it contains evanescent components only. Comparing the same signals crossing either air or a barrier thus makes it possible to measure the signal velocity independently of the preparation and of the detection process.

Superluminal energy velocity became most obvious in a single photon experi-



Fig. 2. Barrier traversal time of a microwave packet through the forbidden band gap of a multilayer structure inside a waveguide (see Fig.1b). The center frequency of the pulse was 8.7 GHz, the pulse width ± 0.5 GHz. The pulses are normalized. The barrier length was 114.2 mm. The velocity of the tunnelled signal was 4.7 *c*. The slow pulse (1) traversed the empty waveguide, whereas the fast one (2) has tunnelled the photonic barrier of the same length [4].

ment carried out by Steinberg *et al.* [3]. In this experiment a photonic lattice barrier was crossed by single photons, and a speed of 1.7c measured.

2. Signals

A signal is a detectable amount of energy that can be used to carry information [5]. Its essential properties will be discussed in this section.

A modern digital signal used in electronical communication is shown in Fig. 3a. The carrier frequency of the signal determines the receiver's address, and the signal half–width represents the information. The signal has been sent 9 000 km along a fiber, and noise is already seen after amplification to the original magnitude, as displayed in the lower part of the figure [6]. A similar single digital signal with a microwave carrier is shown above in Fig. 2. As mentioned above, this signal has traversed a photonic barrier at a superluminal speed of 4.7c.

Fig. 3b shows a mathematical ideal and a frequency band limited sinusoidal signal. The Fourier transform of the frequency band limited signal has non-causal components, *i.e.*, there are already oscillations at negative times. As mentioned above, all signals are frequency band limited. We will discuss the solution of the causality dilemma below.



Fig. 3. (a) Signal used in optical fiber communication. The signal half–width corresponds to the number of bits, *i.e.*, to the transmitted information. The lower one was recorded after a distance of 9 000 km and amplification. Some noise can be seen. The carrier frequency is $2 \cdot 10^{14}$ Hz. The amplitude modulation is limited to a band width of about 10^{10} Hz. (b) Sine wave signal non–frequency (dotted line) and frequency band limited (solid line), frequency is 5 GHz \pm 0.5 GHz. In consequence of the Fourier transform, the frequency band limited signal has signal components at negative times, *i.e.*, before it is switched on.

2.1 Signals are Fundamentally Frequency Band Limited

A single photon can be detected, and delivers information about its energy $\hbar\omega$. Astronomers determine the temperature of a cosmic event by measuring the energy of the emitted photons, for instance in the case of a γ -ray outburst. From the half–width of the photon burst, the total energy involved in the cosmic process can be determined. (Remember, a signal and, in this case, the half–width, are independent

of the signal's magnitude, *i.e.*, the cosmic signal may have travelled either one or a million light years, while the half-width is still the same.)

When analyzing the meaning of the half–width in the present case, we concentrate on amplitude (AM) modulated signals. Let us take, for definiteness, the example of Fig. 3a: the carrier frequency is near $2 \cdot 10^{14}$ Hz, corresponding to the infrared wavelength of 1.5 μ m. The frequency band–width of this AM signal is four orders of magnitude smaller than the carrier frequency.

2.2 Signal Velocity

With the help of Fig. 3a, the signal velocity, *i.e.*, the number of digits, can easily be defined: the complete envelope of the signal has to be detected in order to disentangle the information. The velocity of this envelope defines the signal velocity. (This definition comprises the velocity of the half–width representing the information.) It is only at the end of the signal that the information is obtained, and the desired effect achieved; the velocity of a signal is, loosely speaking, determined by the velocity of both the front and the tail.

Let us now discuss two apparent features of signals that are often addressed in the literature: (i) The information conveyed by a signal is contained in the halfwidth, as elucidated above (in the case of AM). A signal does not depend on its magnitude. If this were the case, any broadcasting station would rapidly face serious problems with increasing distance between receiver and transmitter. The digital signal displayed in Fig. 3a illustrates this point: The half-width does not change as long as the signal's magnitude is above the noise level or the detector's sensitivity. (ii) If frequency band limited signals were given by analytical functions, the complete information would already be contained in the rising edge of the signal. This assumption, extremely difficult to check in the paradigm of a modern signal shown in Fig. 3a, entails strange effects: For instance, when I switch on my office light in the morning, the information about my leaving my office again would be determined at the same time.

Any physical signal has to be frequency band limited. This is a fundamental physical property, as was shown by Nimtz [7]. It is based on Planck's finding that the minimum energy of a field's frequency component is given by $\hbar \omega$ (see Fig. 4). Thus, frequency band unlimited signals containing an infinite frequency spectrum would have infinite energy, contrary to our experience in a finite world.

There is a wellknown dilemma with frequency band limited signals. The Fourier transform of a frequency band limited signal has non-causal forerunners, *i.e.*, Fourier components existing before the signal is switched on [8]. An example of a frequency band limited signal exhibiting these forerunners is shown in Fig. 3b. This problem is solved by the same argument from quantum mechanics used above: the non-causal photon components of a frequency band limited signal are not measurable, since their energy is less than $\hbar\omega$.



Fig. 4. The detectable part of the signal is the part above the straight line representing the limit $E = \hbar \omega$; frequency components below this line, especially the non–causal forerunners, do not have enough energy to be detectable.

2.3 Tunneling Velocity

The superluminal propagation of signals or of single photons with purely evanescent modes measured in different experiments can be adequately described either by the time dependent Schrödinger equation [9, 10] or by the Maxwell equations. This assumption, based on analogies between particle and photonic tunneling [11], has been verified by means of extensive computer simulations [12]. Quite often the argument is given that the ideal mathematical front of a signal travels at the speed c, and cannot be overtaken by the strongly attenuated body of the tunnelled signal. From the mathematical point of view, this is correct. The existence of an ideal front, however, is based on the assumption of an unlimited frequency band required to form the front; such a front necessarily leads to strong signal reshaping as shown in Fig. 5 and discussed in Ref. [13]. A physical signal, even a single photon, is frequency band limited, as discussed above, and the front is not well defined, contrary to the idealised assumptions needed to define a front. Due to the frequency band limitation of a signal and a smooth barrier dispersion relation, no substantial pulse reshaping occurs.

3. Tunneling Time is Universal

Analysis of various experimental data and calculations with different theoretical models point to a universal property of the tunneling process. We have suggested that in general the tunneling time is approximately equal to the reciprocal frequency 1/f of the corresponding tunneling wave packet's frequency [18]. Experimental data from several experimental studies and different photonic barriers are collected in the Table 1. We conjecture that this very universality is valid for all tunneling



Fig. 5. Comparison of normalized intensity vs. time of an airborne signal (solid line) and a tunnelled signal (dotted line) moving from right to left. Both signals have a sharp step at their beginning and the frequency spectrum is infinite. The tunnelled signal is reshaped and attenuated. Moreover, although its maximum has travelled at superluminal speed, both fronts have traversed the same distance with the light velocity c. Here ξ is the maximum of the tunnelled pulse, a is the shift of the maximum, σ is the variance of the tunnelled signal, and σ_0 is the variance of the incoming pulse. It is clearly seen that the latter is longer than the variance of the tunnelled signal.

Photonic Barrier	Reference	Tunneling	Reciprocal
		Time	Frequency
FTIR	Haibel/Nimtz [18]	117 ps	120 ps
at the Double–Prism	Carey et al. [14]	$\approx 1 \text{ps}$	3 ps
	Balcou/Dutriaux [15]	40 fs	11.3 fs
	Mugnai et al. [16]	134 ps	100 ps
Photonic Lattice	Steinberg et al. [3]	1.47 fs	2.3 fs
	Spielmann <i>et al.</i> [17]	2.7 fs	2.7 fs
	Nimtz $et al.$ [4]	81 ps	115 ps
Undersized Waveguide	Enders/Nimtz [1]	130 ps	115 ps

Table 1. Tunneling time data obtained by investigating three types of photonic barriers and measuring at quite different frequencies.

processes, for wave packets either with rest mass or without rest mass.

Data collected from several microwave and optical studies are presented in Table 1. The experiments were carried out with the three different photonic barriers shown in Fig. 1. The conjecture of a universal tunneling time is evident from the data. This surprising property is supported by theoretical data obtained from the Helmholtz and the Schrödinger equations.

4. Summary and Conclusions

Einstein causality, which restricts the velocity of a signal to $v_s \leq c$, is based on the assumption of a frequency band unlimited signal with an ideal front travelling with the speed of light in vacuum, which cannot be overtaken by the body of a signal undergoing pulse reshaping. We have shown that this restriction is violated by evanescent modes: a physical signal whose half-width represents the information is frequency band limited and does not have a well defined front, while the pulse is not markedly reshaped.

We conclude by summarizing the non-classical properties of evanescent modes which were recognized only recently:

- 1) Signals have a finite energy content, and thus, as a consequence of quantum mechanics, a limited frequency band. This is a fundamental physical property [7].
- 2) Tunneling signals may travel at a superluminal speed, including superluminal energy velocity. The superluminal signal reaches the receiver earlier than the airborne signal. This results in a shortened time between cause and effect. However, due to the finite signal length (duration), the past cannot be changed, *i.e.*, the construction of a time machine is not possible.
- 3) A tunneling barrier is traversed in no time. A barrier represents a space without time. (In theology this is called eternity.) Since evanescent modes do not accumulate phase, the predicted phase time velocity, which equals the group and the signal velocities for evanescent modes, is zero. The finite velocity seen in experiments is caused by the interference and resulting phase shift of the incident and reflected wave packets at the barrier's front boundary. In the case of frustrated total internal reflection (FTIR) the tunneling time is due to the Goos–Hänchen shift, as has been shown by Stahlhofen [19, 18]. The phase shift causes a time delay, and thus a finite tunneling time (being independent of barrier length [9, 10].
- 4) Tunneling time is found to be a universal property. The tunneling time measured with opaque barriers equals roughly the reciprocal frequency of the tunneling wave packet [18].

References

- [1] A. Enders and G. Nimtz, J. Phys. 1 (France) 2 (1992) 1693.
- [2] G. Nimtz, Ann. Phys. 7 (1998) 618.
- [3] A Steinberg, P. G. Kwiat, and R. Y. Chiao, Phys. Rev. Lett. 71 (1993) 708.
- [4] G. Nimtz, A. Enders and H. Spieker, J. Phys. 1 (France) 4 (1994) 565.
- [5] National Telecommunication and Information Administration (USA): Fed. Standard 1037C.
- [6] E. Desurvire, Scientific American, 96, January (1992).
- [7] G. Nimtz, Eur. J. Phys. B 7 (1999) 523.
- [8] A. Papoulis, *The Fourier Integral and its Applications*, McGraw-Hill Book Company Inc. New York (1962).
- [9] Th. Hartman, J. Appl. Phys. 33 (1962) 3427.

- [10] A. Enders and G. Nimtz, Phys. Rev. E 48 (1994) 632.
- [11] L. I. Schiff, Am. J. Phys. 25 (1957) 207.
- [12] H. M. Brodowsky, W. Heitmann, and G. Nimtz, Phys. Lett. A 222 (1996) 125.
- [13] Th. Emig, Phys. Rev. E 54 (1996) 5780.
- [14] J. J. Carey, J. Zawadzka, D. Jaroszynski, and K. Wynne, Phys. Rev. Lett. 84 (2000) 1431.
- [15] Ph. Balcou and L. Dutriaux, Phys. Rev. Lett. 78 (1997) 851.
- [16] D. Mugnai, A. Ranfagni, and L. Ronchi, Phys. Letters A 247 (1998) 281.
- [17] Ch. Spielmann, R. Szipcs, A. Stingle, and F. Kraus, Phys. Rev. Lett. 73 (1994) 2308.
- [18] A. Haibel and G. Nimtz, Ann. Phys. (Leizig) 10 (2001) 707.
- [19] A. A. Stahlhofen, Phys. Rev. A 62 (2000) 12112.

Particle Localization and the Notion of Einstein Causality

Gerhard C. Hegerfeldt Institut für Theoretische Physik, Universität Göttingen Bunsenstr. 9, 37073 Göttingen, Germany

The notion of Einstein causality, *i.e.*, the limiting role of the velocity of light in the transmission of signals, is discussed. It is pointed out that Nimtz and coworkers use the notion of signal velocity in a different sense from Einstein, and that their experimental results are in full agreement with Einstein causality in its ordinary sense. We also show that under quite general assumptions instantaneous spreading of particle localization occurs in quantum theory, relativistic or not, with fields or without. We discuss if this affects Einstein causality.

Keywords: superluminal, signal, localization

1. Introduction

The notion of 'Einstein causality' refers to the limiting role of the velocity of light in the transmission of signals. Einstein's principle of finite signal velocity is of fundamental importance for the foundations of physics, both in classical and in quantum physics. If signal velocities could be arbitrarily high, this would either lead to the possibility of absolute clock synchronization and to a change of special relativity, or to the possible existence of superluminal tachyons, with their associated acausal effects [1]. Hence the name Einstein causality.

To be more precise, in this context a *signal* means the experimental creation of any sort of "disturbance" at some space point or small space region and its influence on a measuring device further away. For example, one could produce an electromagnetic pulse and then measure the field strength at some other point. The start time of the signal is the time when the experiment is set into motion, *i.e.*, when the button is pressed. The arrival time of the signal is the *first instance* a measuring device can or does respond. The limiting role of light velocity means that the corresponding time difference divided by the distance cannot exceed *c*.

Nimtz and coworkers [2] have reported superluminal signal velocities in tunneling experiments with microwaves. These experiments and their interpretation, advocated, for example, in the article of Nimtz *et al.* appearing in this issue, has given rise to considerable controversy [3]. It will be shown further below that the controversy is easily resolved by a careful analysis of the notions used by different authors. Nimtz and coworkers employ a definition of signal velocity which is different from the one Einstein had in mind. Using the old definition, it will be seen that the experimental results of Nimtz and coworkers, sophisticated as they are, do not contradict Einstein causality in the original sense, but rather, are in full agreement with it. Thus a conceptual confusion lies at the heart of the matter, which explains a lot of the controversy.

Are there superluminal phenomena in the quantum realm? For a free nonrelativistic particle instantaneous spreading of the wave function is well known. If, at time t = 0, the wave function vanishes outside some finite region V, then the particle is localized in V with probability 1. Instantaneous spreading implies that the probability of finding the particle arbitrarily far away from the initial region is nonzero for any t > 0. In a nonrelativistic theory, however, this superluminal propagation is of no great concern.

If the localization of a free relativistic particle is described by the Newton-Wigner position operator, then instantaneous spreading also occurs, as noted in Refs. [4] and [5] (cf. also Ref. [6]). This also happens for a proposed photon position operator [7]. In 1974 the present author [8] showed that this phenomenon of instantaneous spreading is quite general for a free relativistic particle, irrespective of the particular notion of localization, whether in the sense of Newton-Wigner or some other sense. Later, an alternative proof of this result was given [9], and the result was extended to the center-of-mass motion of relativistic systems with possibly more than one particle [10]. Ruijsenaars and the author [11] then showed that instantaneous spreading occurs for quite general, relativistic or nonrelativistic, interactions. The main result of Ref. [11] was that this instantaneous spreading is mainly due to positivity of the energy plus translation invariance. More recently, it was shown by the author [12] that translation invariance is also not needed. Hilbert space and positivity of the Hamiltonian (energy) suffice to ensure either instantaneous spreading or, alternatively, confinement in a fixed region for all times. Another extension was given by the author [13] for free relativistic particles and for relativistic systems which have exponentially bounded tails in their localization outside some region V. It was shown that the state spreads out to infinity faster than allowed by a probability flow with finite propagation speed. Probably the most astonishing part of our results is the fact that so little is needed to derive them. They hold with and without field theory and with and without relativity. Only Hilbert space and positivity of the energy are needed.

What do these results mean for Einstein causality? This will be discussed in the following, where we concentrate on the role played by positivity of the energy for instantaneous spreading. We also briefly discuss Fermi's two-atom model [14, 15]. But first we turn to the Nimtz controversy

2. Resolution of the Nimtz controversy

Nimtz *et al.* [2] define in Section 2.2 of their paper in this issue what they mean by signal velocity and arrival time. Their definition is motivated by usage in modern engineering. In particular, their notion of arrival time is connected to the read-out time of the signal. However, Einstein had a different meaning in mind when he formulated his principle of the limiting role of the velocity of light for signal velocities, and this has been explained in the Introduction. Definitions are of course neither right nor wrong, but clearly the meaning of a statement as well as its truth depend on the definition of the notions employed in the formulation of the statement. So what do the Nimtz experiments have to say on the question of Einstein causality in its original sense? Are they compatible with it?

In these experiments, typically, a rapid sequence of microwave pulses is generated. Each pulse is split into two and sent over different paths of the same length to a receiver. Calibration of the path length is achieved by displaying the two pulse sequences stroboscopically as still pictures on an oscillograph. Then a photonic tunnel barrier is inserted into one of the paths, which attenuates the corresponding pulses and reshapes them. To compare tunneled and non-tunneled pulses, the former are reamplified to their original amplitude height at the receiving end and again displayed stroboscopically on the oscillograph. The effect is dramatic. Upon insertion of the tunnel barrier, the still picture of the tunneled pulses makes a jump to earlier times, seemingly indicating that they are arriving earlier than the non-tunneled pulses. With the definition of signal velocity and arrival time used by Nimtz and coworkers, this is indeed true.

To see, however, whether this has anything to do with superluminal signal velocities in the Einstein sense, it is astonishing to look at the tunneled pulses without amplification. Experimentally it has been verified by Nimtz and coworkers that the amplitudes of the tunneled pulses are always *below* the amplitude of the nontunneled pulses [22]. In these experiments, the maxima as well as the half widths of the tunneled pulses are ahead of those of the non-tunneled pulses, and therefore arrive earlier. This is graphically depicted in Fig. 1 by the pulses traveling from left to right. The figure is not to scale, and and does not represent experimental curves, but is just for illustration.

For the signal velocity in the Einstein sense, however, the arrival time of the pulse maximum and the read-out time of the half width are not relevant, since they are not used for clock synchronization. What is relevant is the first possible response time of the measuring device, as explained in the Introduction. Now, since experimentally the tunneled pulses are always below the non-tunneled pulses in amplitude, any measuring device will respond first to the non-tunneled pulses and then to the tunneled ones, or at most simultaneously to both. Thus the limiting role of the speed of light G. C. Hegerfeldt



Fig. 1. Typical behavior of airborne pulse (solid line) and tunneled pulse (dashed line), traveling from left to right (not to scale). In the experiments, the amplitude of the latter is always smaller than that of the non-tunneled pulse, although its maximum arrives at an earlier time.

as signal velocity in the sense of Einstein is not violated in the experiments.

What, then, is superluminal here? Let us consider the group and the phase velocity of light. Both are mathematical constructs useful for the description of electromagnetic phenomena. It is well known that both can be larger than c [16], but this cannot be used for superluminal signals in the Einstein sense. Similarly, it has been shown in Ref. [17] that in a somewhat idealized situation the tunneling pulse can be fully described within Maxwell theory by means of another mathematically introduced auxiliary phase-time velocity notion. Again, this auxiliary velocity cannot be used for superluminal signal transmission in the Einstein sense.

So it seems that the controversy about the interpretation of Nimtz's experiments arises from an indiscriminate use of terminology. Terms like signal velocity and arrival time are used by Nimtz and coworkers in a sense different from that of Einstein. When the notions are used in the original sense the experiments are fully compatible with Einstein causality as ordinarily understood.

3. Fermi's two-atom model

To check the speed of light in quantum electrodynamics, Fermi [14] considered two atoms, separated by a distance R and with no photons initially present. One of the atoms was assumed to be in its ground state, the other in an excited state. The latter could then decay with the emission of a photon. Fermi calculated the excitation probability of the atom which had initially been in its ground state. Using standard approximations, he found the excitation probability to be zero for t < R/c. In Ref. [15] the following mathematical result was proved and applied to the Fermi problem. **Theorem:** Let H be a self-adjoint operator, positive or bounded from below, in a Hilbert space \mathcal{H} . For given $\psi_0 \in \mathcal{H}$ let $\psi_t, t \in \mathbb{R}$, be defined as

$$\psi_t = e^{-iHt}\psi_0 \,. \tag{1}$$

Let A be a positive operator in $\mathcal{H}, A \geq 0$, and let $p_A(t)$ be defined as

$$p_A(t) = \langle \psi_t, A\psi_t \rangle . \tag{2}$$

Then either

$$p_A(t) \neq 0$$
 for almost all t (3)

and the set of such t's is dense and open, or

$$p_A(t) \equiv 0 \quad \text{for all} \quad t .$$
 (4)

For the proof, which is based on an analyticity argument, the positivity of both H and A is needed. Positivity means that all expectation values of the operator are nonnegative. Positivity of H alone is not enough. If A is not positive the theorem does not hold. In Eq. (2) one can replace $p_A(t)$ by

$$p_A(t) = \mathrm{tr}Ae^{-iHt}\rho e^{iHt},$$

where ρ is a positive trace-class operator.

If one takes for ψ_0 in the theorem the initial state considered by Fermi and for A the operator describing the excitation probability, *e.g.*, the projector onto the excited states, then $p_A(t)$ becomes the excitation probability, and the theorem states that this probability is immediately nonzero. In [15] it was discussed how to avoid a possible conflict with causality, and this was continued in more detail, for example in [18, 19, 20, 21]. The conclusion was that the immediate excitation could be understood in a field-theoretic context as vacuum fluctuations due to virtual photons. The part of the excitation due to the second atom behaves causally [20, 21]. Causality then holds for expectation values after the spontaneous part has been subtracted. This corresponds to the notion of weak causality, *i.e.*, for expectation values, introduced in [6], which contrasts to the notion of strong causality, *i.e.*, causality for individual events, as discussed in [18]. Fermi seems to have had strong causality in mind.

4. Particle localization and spreading

Let us suppose that it makes sense to speak of the probability of finding a particle at a given time inside a region of space V. This is a highly nontrivial assumption. In a quantum theory the probability of finding a particle or system inside V should be given by the expectation of an operator, N(V), say. Since probabilities lie between 0 and 1, one must have

$$0 \le N(V) \le 1. \tag{5}$$

Now let us assume that the system, with state ψ_0 at t = 0, is strictly localized in a region V_0 , *i.e.*, with probability 1, so that $\langle \psi_0, N(V_0)\psi_0 \rangle = 1$ or, equivalently,

$$\langle \psi_0, (1 - N(V_0))\psi_0 \rangle = 0.$$
 (6)

From Eq. (5) one has $1 - N(V_0) \ge 0$ and hence the theorem can be applied, with

$$A \equiv 1 - N(V_0). \tag{7}$$

As a consequence one either has

$$\langle \psi_l, N(V_0)\psi_l \rangle \equiv 1 \text{ for all } t$$
(8)

or

 $\langle \psi_l, N(V_0)\psi_l \rangle < 1 \quad \text{for almost all } t$. (9)

The alternative in Eq. (8) means that the particle or system stays in V_0 at all times, as might happen for a bound state in an external potential.

Now, if the particle or system is strictly localized in V_0 at t = 0 it is, a fortiori, also strictly localized in any larger region V containing V_0 . If the boundaries of V and V_0 have a finite distance and if finite propagation speed holds, then the probability of finding the system in V would also have to be 1 for sufficiently small times, e.g., $0 \le t < \epsilon$. But then the theorem, with $A \equiv 1 - N(V)$, states that the system stays in V at all times. Now we can make V smaller and smaller and let it approach V_0 . Thus we conclude that if a particle or system is strictly localized in a region V_0 at time t = 0, then finite propagation speed implies that it stays in V_0 at all times, and therefore prohibits motion to infinity. Or put conversely, if there exist particle states which are strictly localized in some finite region at t = 0 and later move towards infinity, then finite propagation speed cannot hold for localization of particles.

This can be formulated somewhat more strongly as follows. If at t = 0 a particle is strictly localized in a bounded region V_0 then, unless it remains in V_0 at all times, it cannot be strictly localized in a bounded region V, however large, for any finite time interval thereafter, and the particle localization immediately develops infinite "tails." The spreading is over all space except possibly for "holes" which, if any, will persist permanently, by the same arguments as before. If the theory is translation invariant, then there can be no holes, as shown in Ref. [11] under some mild spectrum conditions.

5. Counterexample Dirac equation?

At first sight, the Dirac equation might seem to be a counterexample to our results on instantaneous spreading. Indeed, this wave equation is hyperbolic, implying finite propagation speed. For the localization operator N(V) one might take the characteristic function $\chi_V(\mathbf{x})$, just as in the nonrelativistic case and in contrast to the Newton-Wigner operator. Then, for a wave function initially vanishing outside a finite region, *i.e.*, of finite support, the localization does evolve with finite propagation speed! Doesn't this contradict the results of the preceding section?

This example is instructive since it shows the importance of the positive-energy condition. The Dirac equation contains positive and negative energy states. Now, consider a solution of the Dirac equation, which vanishes outside some finite region, and make the additional assumption that it is composed of positive-energy solutions only. Then one gets a contradiction to our results, and therefore the additional assumption must be wrong, *i.e.*, a solution with finite support at some time must contain negative-energy contributions. This means that positive-energy solutions of the Dirac equation always have *infinite* support to begin with! This is phrased as a mathematical result, for instance in the book by Thaller [23].

Thus the results of the preceding section do not apply if there are no strictly localized states in the theory! Strict localization of a state ψ in a region V means that $\langle \psi, N(V)\psi \rangle = 1$, and this gives

$$0 = \langle \psi, (\mathbf{1} - N(V))\psi \rangle = \| (\mathbf{1} - N(V))^{1/2}\psi \|^2,$$

where the root exists by positivity of N(V). This implies

$$N(V)\psi = \psi. \tag{10}$$

Hence ψ is an eigenvector of N(V) for the eigenvalue 1 if ψ is strictly localized in V, and vice versa. The eigenvalue 0 means strict localization outside V.

The existence or nonexistence of strictly localized states depends on the form of N(V). For example, if one has a self-adjoint position operator \hat{X} with commuting components, then N(V) is a projection operator from the spectral decomposition of \hat{X} , and thus has eigenvalues 1 and 0. Hence in this case there are strictly localized states for any region V, and the result of the previous section implies instantaneous spreading.

This instantaneous spreading also occurs for position operators with self-adjoint but *non*-commuting components \hat{X}_i . Each \hat{X}_i has a spectral decomposition whose projection operators give the localization operators for infinite slabs. Eigenvectors for the eigenvalue 1 represent states strictly localized in these slabs, and there is instantaneous spreading in this case, too.

To avoid instantaneous spreading one therefore has to consider localization operators N(V) which are not projectors, for example positive operator-valued measures. However, if one insists on arbitrarily good localization, *i.e.*, on tails which drop off arbitrarily fast, then one runs into our results in Ref. [13].

Discussion. Could instantaneous spreading be used for the transmission of signals if it occurred in the framework of relativistic one-particle quantum mechanics? Let us suppose that at time t = 0 one could prepare an ensemble of strictly localized (non-interacting) particles by laboratory means, *e.g.*, photons in an oven. Then one

could open a window and observe some of them at time $t = \varepsilon$ later on the moon. Or for repetition, suppose one could successively prepare strictly localized individual particles in the laboratory. Preferably this should be done with different, distinguishable, particles in order to be sure when a detected particle was originally released. Such a signaling procedure would have very low efficiency, but could nevertheless be used for synchronization of clocks or, for instance, for betting purposes.

Field-theoretic aspects of our results have been discussed in detail in Ref. [24]. Permanent infinite tails in field theory can be understood intuitively through clouds of virtual particles due to renormalization ('dressed states'). Also, counters could be influenced by vacuum fluctuations.

References

- Cf., e.g., W. Rindler, *Introduction to Special Relativity*, 2nd Edition, Clarendon, Oxford 1991, p. 17. But see also O. M. Bilaniuk, V. K. Deshpande, and E. C. G. Sudarshan, *Am. J. Phys.* 30, 718 (1962).
- [2] G. Nimtz, A. Haibel, A. A. Stahlhofen and R.-M. Vetter, preceding paper in this volume. See also H. Aichmann, A. Haibel, W. Lennartz, G. Nimtz, and A. Spanoudaki, *Proc. of the International Symposium on Quantum Theory and Symmetries*, Goslar, 18-22 July 1999 (ISBN 9-8102-4237-9), 605-611 (2000); G. Nimtz, *Ann. Phys. (Leipzig)* 7, 618 (1998); A. Enders and G. Nimtz, *Phys. Rev. E* 48, 632 (1994), and references therein. For related experiments see R. Y. Ciao and A. M. Steinberg, *Progress in Optics* vol. XXXVII, edited by E. Wolf, Elsevier, Amsterdam (1997), p. 345; L. J. Wang, A. Kuzmich, and A. Dogariu, *Nature* 406, 277 (2000).
- [3] Cf. e.g., H. Goenner, Ann. Phys. (Leipzig) 7, 774 (1998).
- [4] W. Weidlich and K. Mitra, Nuovo Cim. 30, 385 (1963).
- [5] G. N. Fleming, Phys. Rev. B 139, 963 (1965).
- [6] S. Schlieder, in: Quanten und Felder, H.P. Dürr Ed., Vieweg, Braunschweig (1971), p. 145.
- [7] W. O. Amrein, Helv. Phys. Acta 42, 149 (1969).
- [8] G. C. Hegerfeldt, Phys. Rev. D 10, 3320 (1974).
- [9] B. Skagerstam, Int. J. Theor. Phys. 15, 213 (1976).
- [10] J. F. Perez and I. F. Wilde, Phys. Rev. D 16, 315 (1977).
- [11] G. C. Hegerfeldt and S. N. M. Ruijsenaars, Phys. Rev. D 22, 337 (1980).
- [12] G. C. Hegerfeldt, in: *Irreversibility and Causality in Quantum Theory*, edited by A. Bohm, H.-D. Doebner and P. Kielanowski, *Lecture Notes in Physics* 504, p. 238, Springer, Berlin (1998).
- [13] G. C. Hegerfeldt, Phys. Rev. Lett. 54, 2395 (1985).
- [14] E. Fermi, *Rev. Mod. Phys.* 4, 87 (1932); M. I. Shirokov, *Yad. Fiz.* 4, 1077 (1966) [*Sov. J. Nucl. Phys.* 4, 774 (1967)]; W. Heitler and S. T. Ma, *Proc. R. Ir. Acs.* 52, 123 (1949); J. Hamilton, *Proc. Phys. Soc. A* 62, 12 (1949); M. Fierz, *Helv. Phys. Acta* 23, 731 (1950); B. Ferretti, in: *Old and new Problems in Elementary Particles*, edited by G. Puppi, Academic Press, New York (1968), p. 108; P. W. Milonni and P. L. Knight, *Phys. Rev. A* 10, 1096 (1974); M. I. Shirokov, *Sov. Phys. Usp.* 21, 345 (1978); M. H. Rubin, *Phys. Rev. D* 35, 3836 (1987); A. K. Biswas, G. Compagno, G. M. Palma, R. Passante, and R. Persico, *Phys. Rev. A* 42, 4291 (1990); A. Valentini, *Phys. Lett. A* 153, 321 (1991).
- [15] G. C. Hegerfeldt, Phys. Rev. Lett. 72, 596 (1994).
- [16] L. Brillouin, Wave Propagation and Group Velocity, Academic Press, New York (1960).
- [17] Th. Emig, Diplomarbeit, Universität Köln (1995); Phys. Rev 54, 5780 (1996).
- [18] G. C. Hegerfeldt, in: Nonlinear, deformed and irreversible quantum systems, edited by H.-D. Doebner, V.K. Dobrev, and P. Nattermann, World Scientific (1995).
- [19] D. Buchholz and J. Yngvason, Phys. Rev. Lett. 73, 613 (1994).
- [20] A. Labarbara and R. Passante, Phys. Lett. A 206, 1 (1994).
- [21] P. W. Milonni, D. F. V. James, and H. Fearn, Phys. Rev. A 52, 1525 (1995).
- [22] H. Aichmann, private communication, June 1998
- [23] B. Thaller, The Dirac Equation, Springer, Berlin (1992).
- [24] G. C. Hegerfeldt, Ann. Phys. (Leipzig) 7, 716 (1998).

Locality of Quantum Electromagnetic Radiation

Alexander S. Shumovsky Physics Department, Bilkent University, Bilkent, Ankara, 06533, Turkey

We construct the local representation of the Weyl-Heisenberg algebra of multipole photons using the three-dimensional properties of polarization. It is shown that this representation is compatible with the operational approach to photon localization.

Keywords: quantum electrodynamics, electromagnetic radiation, localization

1. Introduction

In spite of the great success of quantum electrodynamics (QED), there remain a number of major unresolved problems (*e.g.*, see [1, 2, 15]). Leaving aside the detailed discussion of foundations of QED, we shall concentrate here on the problem of localization of photons, which has attracted a great deal of interest. The point is that the photon creation and annihilation operators are defined in QED as nonlocal objects. In other words, the photon number operator gives the total number of photons in the volume of quantization without specification of their space-time location [2, 15]. Moreover, it has been proven by Newton and Wigner [16] that no position operator can exist for the photon. There is a widespread belief that the maximum precise localization appears in the form of a wavefront [5]. At the same time, the specific fall-off of the photon energy density and photodetection rate can be interpreted as photon localization in space [6].

Perhaps, the most evident and best example of photon localization is provided by the photodetection process, when a photon is transmitted into an electronic signal in the sensor element of the detecting device [7]. This localization is usually described *operationally* (in terms of what can be measured by a macroscopic detector) by means of the so-called configuration number operator, which determines the number of photons in the cylindrical volume $\sigma e \Delta t$, where σ denotes the area of the sensor element, c is the light velocity, and Δt is detector exposure time [2, 7].

We now stress that, in the usual treatment of photon localization, the radiation field is considered to consist of the plane waves of photons [2, 15]. In reality, the quantum electromagnetic radiation emitted by the atomic and molecular transitions corresponds to *multipole photons* [8] represented by quantized spherical waves [9]. Although the classical plane and spherical waves are equivalent in the sense that

they both form complete orthogonal sets of solutions of the homogeneous Helmholtz wave equation [10], there is a strong qualitative difference between the two quantum representations. The plane waves of photons correspond to the running-wave solution in empty space with translational symmetry, which leads to states of photons with given linear momentum. In turn, the solution in terms of spherical waves assumes the existence of a singular point, corresponding to an atom (source or absorber of radiation) whose size is small with respect to the wavelength. In this case, the boundary conditions correspond to the rotational symmetry, and lead to states of photons with given angular momentum. Since the components of linear and angular momenta do not commute, the two representations of the quantum electromagnetic field correspond to physical quantities which cannot be measured at the same time.

The main objective of this paper is to show that the use of the multipole photon representation leads to an adequate description of localization in the atom-field interaction process. The paper is arranged as follows. In Section 2 we briefly discuss the difference between the spatial properties of plane and spherical waves of photons. In Section 3 we introduce the local representation of the multipole photon. Then, in Section 4, we discuss the problem of measurement and causality. A general conclusion and the implications of this work are presented in Section 5.

2. Plane and spherical waves of photons

An arbitrary free quantum electromagnetic field can be described by the operator vector potential whose positive-frequency part has the following form

$$\vec{A}^{(+)}(\vec{r},t) = \sum_{\mu=-1}^{1} (-1)^{\mu} \vec{\chi}_{-\mu} \sum_{k,\ell} V_{k\ell\mu}(\vec{r}) e^{-i\omega_k t} a_{k\ell}, \tag{1}$$

where the unit vectors

$$\vec{\chi}_{\pm} = \mp \frac{\vec{e}_x \pm i \vec{e}_y}{\sqrt{2}}, \qquad \vec{\chi}_0 = \vec{e}_z,$$
(2)

form the so-called helicity or spin basis of the three-dimensional space [9, 11], $V_{k\ell\mu}(\vec{r})$ is the mode function, and $a_{k\ell}$ is the photon annihilation operator, which obey Weyl-Heisenberg commutation relations

$$[a_{k\ell}, a^+_{k'\ell'}] = \delta_{kk'} \delta_{\ell\ell'}.$$
(3)

Here ℓ is a cumulative index. By construction, the vector potential components $A_{\mu=\pm 1}(\vec{r},t)$ in (1) describe the circularly polarized transversal components of the field with positive and negative helicity respectively, while $A_{\mu=0}(\vec{r},t)$ gives the linearly polarized longitudinal component [11]. In the case of plane waves of photons

$$\vec{\chi}_0 = \frac{\vec{k}}{k},$$

and projection of spin of the photon on this axis is forbidden, so that there are only two transversal components of the field. In this case, index $\ell \equiv \sigma = \pm$ describes the circular polarization of the field.

Unlike the plane waves of photons, the quantum multipole radiation has all three spatial components [9, 11], and index $\ell = \{\lambda, j, m\}$ gives the parity $\lambda = E, M$ (type of radiation, either electric or magnetic), angular momentum of photons $j = 1, 2, \cdots$, and projection of the angular momentum on the quantization axis $m = -j, \cdots, j$. It should be stressed that plane and multipole photons have different numbers of quantum degrees of freedom. In fact, a monochromatic radiation field has only two degrees of freedom, described by the polarization index $\sigma = \pm$ in the case of plane photons. At the same time, a monochromatic multipole field of a given type λ at given $j \ge 1$ is specified by $(2j+1) \ge 3$ degrees of freedom. Moreover, the polarization is not a quantum number and, thus, the global property of the multipole radiation changes from point to point [12].

The spatial properties of the field are described by the mode functions in (1). In the case of plane photons, the mode function has the simple form of plane waves (e.g., see [2])

$$V_{k\sigma}(\vec{r}) = \gamma e^{i\vec{k}\cdot\vec{r}}, \qquad \gamma = \sqrt{\frac{2\pi\hbar c}{kV}},\tag{4}$$

where V is the volume of quantization. It is seen that this expression leads to the spatially homogeneous density of intensity of a monochromatic plane wave

$$I^{(plane)} = \vec{E}^{(-)}(\vec{r}) \cdot \vec{E}^{(+)}(\vec{r})$$

= $k^2 \vec{A}^{(-)}(\vec{r}) \cdot \vec{A}^{(+)}(\vec{r}) = (k\gamma)^2 \sum_{\sigma} a_{\sigma}^+ a_{\sigma}.$ (5)

In turn, the multipole radiation is specified by the mode functions [9, 11]

$$V_{Ekjm\mu} = \gamma_{Ekj} [\sqrt{j} f_{j+1}(kr) \langle 1, j+1, \mu, m-\mu | jm \rangle Y_{j+1,m-\mu}(\theta\phi) - \sqrt{j+1} f_{j-1}(kr) \langle 1, j-1, \mu, m-\mu | jm \rangle Y_{j-1,m-\mu}(\theta,\phi)], V_{Mkjm\mu} = \gamma_{Mkj} f_j(kr) \langle 1, j, \mu, m-\mu | jm \rangle Y_{jm}(\theta,\phi)$$
(6)

in the case of $\lambda = E$ and $\lambda = M$, respectively. Here $\langle \cdots | jm \rangle$ denotes the Clebsch-Gordon coefficient of vector addition of spin and orbital parts of the angular momentum, $Y_{\ell n}$ is the spherical harmonics, and

$$\gamma_{\lambda} = \begin{cases} \gamma/\sqrt{2j+1}, & \text{at } \lambda = E\\ \gamma & \text{at } \lambda = M. \end{cases}$$

The radial dependence in (6) is defined as follows [10]

(. . .

$$f_{\ell}(kr) = \begin{cases} h_{\ell}^{(1)}(kr), & \text{outgoing spherical wave} \\ h_{\ell}^{(2)}(kr), & \text{incoming spherical wave} \\ j_{\ell}(kr), & \text{standing spherical wave,} \end{cases}$$
(7)

where $h_{\ell}^{(1,2)}$ denotes the spherical Hankel function of the first and second kind respectively and j_{ℓ} is the spherical Bessel function. It is clear that, unlike (5), the density of intensity of a monochromatic pure *j*-pole multipole radiation of a given type

$$I^{(multi)}(\vec{r}) = \sum_{\mu} \sum_{m,m'=-j}^{j} V^*_{\lambda k j m \mu}(\vec{r}) V_{\lambda k j m' \mu}(\vec{r}) a^+_{\lambda k j m} a_{\lambda k j m'}$$
(8)

shows a certain position dependence with respect to the source location at the origin of the reference frame spanned by the helicity basis (2). This spatial inhomogeneity of the density of intensity of multipole radiation can be used to introduce the local representation of the Weyl-Heisenberg algebra of multipole photons [12].

3. Local photon operators

In contrast to (5), the density of intensity (8) is represented by a non-diagonal form in the photon operators which can be represented as follows

$$I^{(multi)}(\vec{r}) = \sum_{m,m'} \mathcal{V}_{mm'}(\vec{r}) a_m^+ a_m',$$
(9)

where $\mathcal{V}(\vec{r})$ is the Hermitian $(2j+1) \times (2j+1)$ matrix with the elements

$$\mathcal{V}_{mm'}(\vec{r}) = k^2 \sum_{\mu=-1}^{1} V_{m\mu}^*(\vec{r}) V_{m'\mu}(\vec{r}).$$
(10)

To simplify the notations, hereafter we omit the indexes λ , k, and j. It is seen that

$$tr\mathcal{V}(\vec{r}) \equiv \sum_{m} \mathcal{V}_{mm}(\vec{r}) = k^2 [\vec{A}^{(+)}(\vec{r}), \vec{A}^{(-)}(\vec{r})], \tag{11}$$

so that the trace of (10) describes the electric-field contribution into the energy density of the zero-point oscillations [14] of the multipole field. Then

$$W_{\mu}(\vec{r}) = k^2 \sum_{m=-j}^{j} |V_{m\mu}(\vec{r})|^2$$
(12)

gives the contribution of spatial components with different polarization μ into the zero-point energy of the multipole field. Since the polarization is the three-dimensional property of the multipole radiation [13, 14], it seems to be reasonable to define the spatial properties of multipole photons by means of polarization.

Consider for definiteness the electric type pure j-pole monochromatic radiation. Then, the operator polarization matrix takes the form [13]

$$P_{\mu\mu'}(\vec{r}) = E^{(-)}_{\mu}(\vec{r})E^{(+)}_{\mu'}(\vec{r}) = k^2 A^{(-)}_{\mu}(\vec{r})A^{(+)}_{\mu'}(\vec{r}).$$
(13)

By definition, this is the (3×3) Hermitian matrix with the operator elements written in the normal order. In addition, one can define the anti-normal operator polarization matrix. Then, the difference between the anti-normal and normal matrices defines the zero-point oscillations of polarization [15] with the elements

$$P_{\mu\mu'}^{(0)}(\vec{r}) = k^2 [A_{\mu}^{(+)}(\vec{r}), A_{\mu'}^{(-)}(\vec{r})] = k^2 \sum_m V_{m\mu}(\vec{r}) V_{m\mu'}^*(\vec{r}).$$
(14)

It is easily seen that the diagonal elements of (14) coincide with (12). It is intuitively clear that the spatial properties of the zero-point oscillations of polarization described by (14) should be determined by distance r from the source *independent of the spherical angles*. In other words, the vacuum noise should have a homogeneous angular distribution, which can change with the distance.

The (3×3) Hermitian matrix (14) can be diagonalized by a proper transformation of the reference frame spanned by the helicity basis (2)

$$U(\vec{r})P^{(0)}(\vec{r})U^{+}(\vec{r}) = \mathcal{P}^{(0)}(\vec{r}), \qquad U^{+}(\vec{r})U(\vec{r}) = \mathbf{1}.$$
(15)

As a result of this transformation,

$$\vec{\chi}_0 \rightarrow \vec{\chi}'_0 = \vec{r}/r.$$

It is then a straightforward matter to arrive at the conclusion that

$$\mathcal{P}^{(0)}(r) = \begin{pmatrix} P_T(r) & 0 & 0\\ 0 & P_L(r) & 0\\ 0 & 0 & P_T(r) \end{pmatrix},$$
(16)

where

$$P_T(r) = k^2 |V_{\mu\mu}(\vec{r})|^2, \text{ at } \mu = \pm 1,$$

$$P_L(r) = k^2 |V_{\mu\mu}(\vec{r})|^2, \text{ at } \mu = 0.$$
(17)

In other words, the diagonal elements in (16) describe the transversal and longitudinal (with respect to $\vec{\chi}'_0$) vacuum noise of polarization as a function of distance from the source.

The use of the same unitary transformation (15) allows the operator polarization matrix (13) to be cast in the form

$$\mathcal{P}(\vec{r}) = U(\vec{r})P(\vec{r})U^+(\vec{r}), \qquad (18)$$

where

$$\mathcal{P}^{(E,n)}_{\mu\mu'}(\vec{r}) = k^2 \mathcal{A}^+_{Ekj\mu}(\vec{r}) \mathcal{A}_{Ekj\mu'}(\vec{r}), \qquad (19)$$

and

$$\mathcal{A}_{Ekj\mu}(\vec{r}) = \sum_{\mu'=-1}^{1} U^*_{\mu\mu'}(\vec{r}) \sum_{m=-j}^{j} V_{Ekjm\mu'}(\vec{r}) a_{Ekjm}.$$
 (20)

In view of (3), the operators (20) obey the commutation relations

$$\left[\mathcal{A}_{\lambda k j \mu}(\vec{r}), \mathcal{A}^{+}_{\lambda' k' j' \mu'}(\vec{r})\right] = \delta_{\lambda \lambda'} \delta_{k k'} \delta_{j j'} \delta_{\mu \mu'} \times \begin{cases} P_T(r) & \text{at } \mu = \pm 1\\ P_L(r) & \text{at } \mu = 0 \end{cases}$$
(21)

where P_L , P_T are the diagonal elements (17). Similar results can also be obtained in the case of the magnetic multipole radiation.

We now note that the only difference between (3) and (21) is the presence of position-dependent factors in the right-hand side of (21). It seems to be tempting to introduce the normalized local operators

$$b_{\lambda k j \mu}(\vec{r}) = \frac{\mathcal{A}_{\lambda k j \mu}(\vec{r})}{\sqrt{P_{\mu}(\vec{r})}},$$
(22)

which obey the standard Weyl-Heisenberg commutation relations

$$[b_{\lambda k j \mu}(\vec{r}), b^{+}_{\lambda' k' j' \mu'}(\vec{r})] = \delta_{\lambda \lambda'} \delta_{k k'} \delta_{j j'} \delta_{\mu \mu'}$$
(23)

at any point \vec{r} . Hence, the transformation (15) can be interpreted as a local Bogolubov canonical transformation [17], conserving the Weyl-Heisenberg commutation relations. In fact, the equations (15) and (22) describe the transformation of global multipole photon operators $a_{\lambda kjm}$ with given $m = -j, \dots, j, j \ge 1$, into the local photon operators $b_{\lambda kj\mu}(\vec{r})$ with given polarization μ at any point of the space.

4. Measurement and locality

In the operational approach to photon localization [7] (also see [2, 15]), the local absorption operator

$$\vec{a}(\vec{r},t) = \gamma \sum_{k,\sigma}' e^{i(\vec{k}\cdot\vec{r}-kct)} \vec{e}_{k\sigma} a_{k\sigma}$$
(24)

is defined in the case of plane waves of photons. Here summation is taken over a finite set of modes to which a detector responds. Then, the so-called *configuration space number operator* is defined by the relation

$$\mathcal{N}(\mathcal{V},t) = \int' \vec{a}^{+}(\vec{r},t) \cdot \vec{a}(\vec{r},t) d^{3}r$$
$$= \gamma^{2} \sum_{k,\sigma} \sum_{k',\sigma'} \vec{e}_{k\sigma} \cdot \vec{e}_{k'\sigma'} e^{-i(\vec{k}-\vec{k}')\cdot\vec{r})} e^{i(k-k')ct} a^{+}_{k\sigma} a_{k'\sigma'}, \tag{25}$$

where the integral is taken over the volume of photon localization (cylinder with base corresponding to the sensitive area of the detector and height proportional to the exposure time). It is clear that the operators (24) and (25) obey the following commutation relations

$$[\mathcal{N}(\mathcal{V}, t), \mathcal{N}(\mathcal{V}', t)] = 0 \tag{26}$$

and

$$[\vec{a}(\vec{r},t),\mathcal{N}(\mathcal{V},t)] \approx \begin{cases} -\vec{a}(\vec{r},t), & \text{if } \vec{r} \in \mathcal{V} \\ 0, & \text{otherwise,} \end{cases}$$
(27)

where \mathcal{V} denotes the volume of localization (detection). Let us stress that (27) has an approximate sense.

There is a principal difference which makes difficult the direct use of the operational approach to the problem of localizing photons in the case of multipole radiation. The point is that the multipole photons are in the state with given angular momentum, and therefore they have no well defined direction of propagation. At the same time, these photons are localized initially inside the source.

Let us now note that the operators (22) describe the local properties of the multipole radiation, and that the density of intensity operator (9) can be represented by

$$I^{(multi)}(\vec{r}) = \sum_{\mu=-1}^{1} b_{\mu}^{+}(\vec{r}) b_{\mu}(\vec{r}).$$
(28)

Under the condition that the we have a strictly monochromatic field, the operator (28) can be considered as an analog of (25) at a given point \vec{r} , while (22) is similar to (24). The principal difference between the two local representations is that

$$[b_{\lambda k j \mu}(\vec{r}), b^{+}_{\lambda' k' j' \mu'}(\vec{r}')] = \delta_{\lambda \lambda'} \delta_{k k'} \delta_{j j'} f_{\mu \mu'}(\vec{r}, \vec{r}'), \qquad (29)$$

where $f_{\mu\mu'}(\vec{r}, \vec{r'})$ is, perhaps, a sharp function but $f_{\mu\mu'}(\vec{r}, \vec{r'}) \neq \delta_{\mu\mu'}\delta(\vec{r} - \vec{r'})$. Such a violation of the Weyl-Heisenberg commutation relations reflects the causal dependence between the multipole radiation fields at different points.

Nevertheless, the operators (22) can be used for description of a real measurement. Consider a model of a Hertz-type experiment on emission and detection of multipole photons in the system of two identical atoms separated by a distance *d*. If we assume that a photon is first emitted by the atom number one (source) and then absorbed by the atom number two (detector), it is most natural to consider the field as a superposition of outgoing and incoming spherical waves focused on the source and detector respectively. This superposition should obey the boundary conditions for the real radiation field, so that only one multipole photon can exist in the space. Then, in direct analogy to (24), we can construct a configuration space photon absorption operator

$$\vec{\mathbf{a}}(\vec{r},t) = \sum_{\mu} (-1)^{\mu} \vec{\chi}_{-\mu} \sum_{\lambda k j m}^{\prime} V_{\lambda k j m \mu}(\vec{r}) e^{-ikc} a_{\lambda k j m}.$$
 (30)

Here the sum is taken over the modes allowed by the selection rules for the atomfield interaction under consideration. The volume of detection is defined in this case as follows

$$\mathcal{V} = \frac{4\pi}{3} [(c\Delta\tau)^3 - r_a^3],$$

where $\Delta \tau$ is the atomic "exposure time" defined by the natural breadth of the spectral line, and r_a denotes the atomic radius. Then, the configuration space multipole photon number operator takes the form

$$\mathcal{N}(\mathcal{V},t) = \sum_{\mu} \int_{\mathcal{V}} b_{\mu}^{+}(\vec{r},t) b_{\mu}(\vec{r},t) P_{\mu}(\vec{r}) d^{3}r,$$
(31)

where the definition of $b_{\mu}(\vec{r},t)$ differs from (22) by summation over all allowed modes, which induces the time dependence. It is straightforward to show that the operators (30) and (31) obey commutation relations of the type (26) and (27). Thus, the picture of measurement in the source-detector system of two identical atoms expressed in terms of the local operators (22) is compatible with Mandel's operational approach to the photon localization.

The above picture, based on the superposed state of outgoing and incoming waves of photons, completely eliminates an enquiry concerning the trajectory of photons between the atoms. In fact, the quantum mechanical path of a photon is not a well-defined notion [17]. The most that we can state about the path of a quantum particle in many cases is that it is represented by a nondifferentiable, statistically self-similar curve [17]. For example, the path of a tunneling electron and time spent in the barrier are not still defined unambiguously [18]. Moreover, recent experiments on photonic tunneling and transmission information show the possibility of superluminal motion of photons inside an opaque barrier [19].

We now note that, according to the principles of quantum theory, not the path, but causality in the transmission of information from one object to another, is important. In the above considered Hertz experiment with two atoms separated by empty space, this means that the detecting atom cannot be excited earlier than d/c seconds after the emission of a photon by the first atom. Such a causality has been proven in [20].

5. Conclusion

Let us briefly discuss the results obtained here. It has been shown that the clear-cut distinction between the properties of plane and spherical waves leads to a qualitative difference in the spatial behaviour of the corresponding photons as well as of the zero-point oscillations. The successive use of the spatial inhomogeneity of multipole radiation permits us to construct a local representation of the Weyl-Heisenberg algebra of multipole photons based on the properties of polarization. Let us stress once more that the polarization defined to be the spin state of photons has a one-to-one correspondence with the spatial properties of radiation. The local representation

of multipole photons obtained in Sec. 3 is compatible with Mandel's operational definition of photon localization [7]. It permits us to describe a complete Hertz-type experiment with two identical atoms used as the emitter and detector. The two-atom Hertz experiment can be realized for the trapped Rydberg atoms [21]. Finally, we stress the fact that this measurement is closely connected with the problem of engineered atomic entanglement discussed in [22].

References

- [1] E.A. Power, Introductory Quantum Electrodynamics (Longman, London, 1964);
 - I. Bialynicki-Birula and Z. Bialynicka-Birula, *Quantum Electrodynamics* (Pergamon Press, Oxford, 1976);
 - H.A. Kramers, Quantum Mechanics (North-Holland, Amsterdam, 1985);
 - M.W. Evans and J.-P. Vigier, The Enigmatic Photon, Vol. 1 and 2 (Kluwer, Dordreht, 1994);
 - M.W. Evans, J.-P. Vigier, S. Roy, S. Jeffers, and G. Hunter, *The Enigmatic Photon*, Vol. 3 and 4 (Kluwer, Dordreht, 1997).
- [2] L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics* (Cambridge University Press, New York, 1995).
- [3] M.O. Scully and M.S. Zubairy, Quantum Optics (Cambridge University Press, New York, 1997).
- [4] T.D. Newton and E.P. Wigner, Rev. Mod. Phys. 21, 400 (1949).
- [5] R. Acharya and E.C.G. Sudarshan, J. Math. Phys. 1, 532 (1960).
- [6] I. Bialynicki-Birula, Acta Phys. Pol. A 86, 97 (1994);
 I. Bialynicki-Birula, Phys. Rev. Lett. 80, 5247 (1998).
- [7] L. Mandel, Phys. Rev. 144, 1071 (1966).
- [8] E.U. Condon and G.H. Shortley, *The Theory of Atomic Spectra* (Cambridge University Press, New York, 1987).
- W. Heitler, *The Quantum Theory of Radiation* (Oxford University Press, New York, 1954);
 C. Cohen-Tannouji, J. Dupont-Roc, and G. Grinberg, *Atom-Photon Interaction* (Wiley, New York, 1992).
- [10] H. Bateman, *The Mathematical Analysis of Electric and Optical Wave-Motion* (Dover, New York, 1955):

- [11] A.S. Davydov, Quantum Mechanics (Pergamon Press, Oxford, 1976).
- [12] A.S. Shumovsky, in Electronic Proc. of 7th Central European Workshop on Quantum Optics, April 28 May J, 2000 http://optics.szfki.kfki.hu/ cewqo2000.
- [13] A.S. Shumovsky and Ö.E. Müstecaplioğlu, Phys. Rev. Lett. 80, 1202 (1998).
- [14] A.S. Shumovsky, J. Phys. A 32, 6589 (1999).
- [15] W.H. Louisell, Radiation and Noise in Quantum Electronics (McGraw-Hill, New York, 1964).
- [16] N.N. Bogolubov Selected Works, Part II, Quantum and Classical Statistical Mechanics, edited by N.N. Bogolubov, A.M. Kurbatov, and A.S. Shumovsky (Gordon and Breach, New York, 1991).
- [17] R.P Feynman and A.R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).
- [18] E.H. Haug and J.A. Støvneng, *Rev. Mod. Phys.* 61, 917 (1989);
 R. Landauer, *Nature* 341, 567 (1989);
 R. Landauer and Th. Martin, *Rev. Mod. Phys.* 66, 217 (1994).
- [19] A. Enders and G. Nimtz, J. Phys. I(France) 2, 1693 (1992);
 A. Ranfagni, P. Fabeni, G. Pazzi, and D. Mugnai, Phys. Rev. E 48, 1453 (1993);
 Ch. Spielmann, R. Szipöcs, A. Stingl, and F. Kraus, Phys. Rev. Lett. 73, 2308 (1994);
 D. Mugnai, A. Ranfagni, and R. Ruggeri, Phys. Rev. Lett. 84, 4830 (2000).
- [20] D. Kaup and V.I. Rupasov, J. Phys. A 29, 6911 (1996).
- [21] S. Haroche, AIP Conf. Proc. Vol. 464, Issue 1, p. 45 (1999).
- [22] M.B. Plenio, S.F. Huelga, A. Beige, and P.L. Knight, Phys. Rev. A 59, 2468 (1999).

J.D. Jackson, Classical Electrodynamics (Wiley, New York, 1978).

Time of Arrival in Classical and Quantum Mechanics

Juan León Instituto de Matemáticas y Física Fundamental, CSIC Serrano 113-bis, E-28006 MADRID, Spain

The time of arrival at an arbitrary position in configuration space can be given as a function of the phase space variables for the Liouville integrable systems of classical mechanics, but only for them. We review the Jacobi-Lie transformation that explicitly implements this function of positions and momenta. We then discuss the recently developed quantum formalism for the time of arrival. We first analyze the case of free particles in one and three space dimensions. Then, we apply the quantum version of the Jacobi-Lie transformation to work out the time of arrival operator in the presence of interactions. We discuss the formalism and its interpretation. We finish by disclosing the presence (absence) of "instantaneous" tunneling for thin (thick) barriers.

Keywords: time; phase space; Hilbert space; positive operator valued measures; tunneling

1. Introduction

Classical and Quantum Mechanics use the notion of Newtonian time, a universal parameter that rules the evolution of all the dynamical systems of the universe. Newtonian time is "a priori" external to everything, physical systems and observers alike. However, in many instances there are true time-like properties in the physical systems under study. In general, the answers to questions like: *How long will it take to ...?* or, *When will it ...?, etc.* come in the form of a time that genuinely depends of the very system. The crux of the matter is finding the time (the time elapsed, or the instant in time) in which some property of the system will take a specific value, something that could be generically termed as "the time of arrival at that value." In the next section we deal with the formulation of this question in classical dynamics. The much more involved case of translating a time parameter into an operator on the Hilbert space, as required by the quantum treatment, is worked out in sections 2 and 3 for free and interacting particles respectively. In section 4 we point out some eccentric properties of the time of arrival at places that are classically forbidden.

2. Deriving time in phase space

The treatment of time as a phase space variable is a time-honored procedure. The term *extended phase space* was coined for the approach in which, to the n pairs (q, p) of the phase space variables of mechanical systems with n degrees of freedom,
one adds the additional conjugate pair (t, p_t) , which requires the constraint $p_t + H(q, p) = 0$ for consistency. It seems possible to dismantle this construction by replacing the pair (t, p_t) by another canonical pair (q, p). Naively, one would single out one of the phase space variables $(q_1$ for instance) and make it equal to some parametric value (i.e. $q_1 = x$). Then, its canonical conjugate (the momentum p_1 in this example), would be fixed by the constraint, giving $p_1 = \phi(x; q_2, p_2, \ldots, p_t)$. The phase space would now be given in terms of $\{(q_2, p_2), \ldots, (q_n, p_n), (t, p_t)\}$, with x acting as an external evolution-like parameter. Hence, t(x) or in words, "the time of arrival at x" would be a legitimate question to ask. In spite of its apparent generality, it is seldom possible to accomplish this program, not because of its very difficulty, but due to the non-fulfillment of some of the many conditions necessary for the existence of solutions. Here we will discuss the case of integrable systems for which there is a global construct for t(x), that we will describe explicitly.

In the modern approach to classical dynamics (a standard reference is [1], a very readable text can be found in [2]), a Hamiltonian system is called completely integrable (*a là* Liouville) when it satisfies the conditions a and b below:

a.2.
$$\{\Phi_i, \Phi_j\} = 0, \forall i, j = 1, \dots, n.$$

b. The conservation laws define n isolating integrals that can be written as:

b.1.
$$\Phi_i = C_i \Rightarrow p_i = \phi_i(q_1, \dots, q_n, C_1, \dots, C_n; t),$$

 $\forall i = 1, \dots, n.$
b.2. $\frac{\partial \phi_i}{\partial q_j} = \frac{\partial \phi_j}{\partial q_i} \quad \forall i, j = 1, \dots, n.$

In these conditions, the solution to the Hamilton equations is an integrable flow, described by a system of holonomic coordinates (q(t), p(t)) in phase space for each instant of time:

$$q_i(t) = q_i(q_0, p_0; t), \quad i = 1, \dots, n.$$
(1)
$$p_i(t) = p_i(q_0, p_0; t), \quad i = 1, \dots, n.$$

In particular, given the set of initial conditions (q_0, p_0) , the system arrives at the point (q(t), p(t)) of the phase space in the path independent instant t. Conversely, these points define the corresponding times of arrival. In this case, time meets the requirements to qualify as a derived variable in phase space, whose explicit construction occupies the rest of this section.

For integrable flows there is a special choice of phase space coordinates that mathematically eliminates the effects of interactions (because the new positions are ignorable coordinates). In other words, integrable systems are canonically equivalent to a set of translations (or of circular motions) at constant speed. It is customary to denote the variables that determine these translations as action-angle variables, a name which strictly is appropriate only for the case of periodic systems, where the (closed) flow lines are topologically equivalent to circles. For these integrable flows, there is a canonical transformation W (the Jacobi-Lie transformation) to free-like variables

$$\{q, p; H(q, p)\} \xrightarrow{W} \{Q, P; H_0(P)\},$$
 (2)

where $H(q, p) = H_0(P)$. The most useful form of this transformation is W(q, P), that is, a function of the old positions and the new momenta, so that

$$Q_i = \frac{\partial W(q, P)}{\partial P_i}, \quad p_i = \frac{\partial W(q, P)}{\partial q_i} \quad i = 1, \dots, n,$$
(3)

The choice $H_0(P) = \sum_i \frac{P_i^2}{2m}$ relates the free coordinates $P_i(t) = P_i$ and $Q_i(t) = \frac{P_i}{m}t + Q_i$ of the translation flow to the positions and momenta $(q_i(t), p_i(t))$ of the actual flow generated by $H(q, p) = \sum_i \frac{p_i^2}{2m} + V(q)$. In this work we shall only consider unbound systems with positive energy $H = H_0 \ge 0$. For this reason we choose, as constant variables, the conserved momenta P_i instead of the usual actions over a period $\oint pdq$ that are more apt for bounded motions. Notice however that the P_i are different from the momenta appearing in perturbative calculations, even if both sets may coincide asymptotically or in some set of R^n . Coming back to our problem, the function W would be given explicitly as a complete integral of the following Hamilton-Jacobi equation:

$$H(q_i, \frac{\partial W(q, P)}{\partial q_i}) = \sum_{i=1}^n \frac{P_i^2}{2m}.$$
(4)

Due to the relations b.1 and b.2 above, it is permitted to write W as the path-independent integral:

$$W(q,P) = \sum_{i=1}^{n} \int_{q_0}^{q} dq_i \phi_i(q,C),$$
(5)

where q_0 is a constant configuration space point, and the C_i (that remain fixed during the integration) are functions of the P_i whose determination is necessary to solve the problem explicitly. We are not concerned here with the search for specific solutions, but with the fact that integrability ensures their global existence. In fact, the equations (3) can be written in the form

$$p_a(q,P) = \phi_a(q,C), \quad Q_a(q,P) = \sum_{i=1}^n \int_{q_0}^q dq_i \frac{\partial p_i(q,P)}{\partial P_a}, \quad a = 1, \dots, n.$$
(6)

J. Leon

The first equation is simply the definition of the isolating integrals (b.1). As a bonus, time can be given as a function of phase space in two alternative ways: either in terms of the old variables, or equally in terms of the new ones. Consider that a particle initially at (\mathbf{q}, \mathbf{p}) arrives at the position $\mathbf{q}(t) = \mathbf{x}$ in the instant $t(\mathbf{x}) = t$, then:

$$t(\mathbf{x}) = \frac{m}{P_a} (X_a - Q_a) = \frac{m}{P_a} \sum_{i=1}^n \int_{\mathbf{q}}^{\mathbf{x}} dq_i \frac{\partial p_i(\mathbf{q}, \mathbf{P})}{\partial P_a}, \quad a = 1, \dots, n,$$
(7)

where $X_a = \partial W(\mathbf{x}, \mathbf{P}) / \partial P_a$ (obviously, $X_a = Q_a(t(\mathbf{x}))$ by construction). Note that in (7) there is no summation over the index a. In fact, integrability can be envisioned as the simultaneous existence of n independent flows, each of them contained in a different phase space plane. The requirement of integrability was noticed by Einstein [3], who analyzed its implications for the old Bohr-Sommerfeld quantization conditions, which he reformulated accordingly giving a new condition that was criticized by Epstein [4]. Integrability [5] allows n different expressions to define the unique time of arrival. Only a pair (Q_a, P_a) appears in each of them, and they all are equivalent. This holds even when there is no separable solution to the original Hamilton-Jacobi equation (4) due to the presence of the potential $V(\mathbf{q})$ in the Hamiltonian. Only for some well known cases [6] the problem is separable in the original variables. Independently of this, notice that as $(\mathbf{Q}(t), \mathbf{P})$ defines a straight line in the phase space, it is simple to lay one of the axes (the n^{th} say) along it. This amounts to defining $H_0(\mathbf{P}) = P_n$, which gives $P_n(t) = E$ and $Q_n(t) = t + Q_n$, while the other variables remain constant $Q_i(t) = Q_i, P_i(t) = P_i, j = 1, ..., n-1$. With this choice, one can write:

$$t(\mathbf{x}) = \sum_{i=1}^{n} \int_{q}^{x} dq_{i} \frac{\partial p_{i}(q_{1} \dots q_{n}, P_{1} \dots P_{n-1}, E)}{\partial E},$$
(8)

with the p_i 's given in (6). This is the standard equation of time that appears in the literature. The rest of the relations would give the time independent geometric properties of the trajectories. Note that for central potentials only p_r depends on E, so that (8) reduces to $t(r_x) = \int_{r_0}^{r_x} dr(\partial p_r/\partial E)$.

We have focused the discussion of this section on the dual definition of the time of arrival, that can be given in terms of the original phase space variables, or of the free translation variables. This duality is a foundation stone for the quantum method presented in this paper. We will obtain the time of arrival operator of interacting particles $\hat{t}(x)$ by applying a quantum version of the canonical transformation W(q, P)to the well known operator for the time of arrival of free particles $\hat{t}_0(x)$. The properties of the latter have been extensively analyzed in the literature. For completeness, and to fix the notation, we present a summary of them in the next section.

3. Time of arrival of free quantum particles

In one space dimension Eq. (7) gives the time of arrival at x of a free particle initially at (q, p) as a function of the phase space variables that depends on x parametrically: $t_0(q, p; x) = m(x - q)/p$. In spite of its simplicity, this expression presents serious quantization difficulties, [7, 8, 9, 10, 11], whose solution we outline here [10, 12, 13, 14, 15]. First of all, it requires a decision about operator ordering, the simplest one being symmetrization:

$$\hat{t}_0(\hat{q}, \hat{p}; x) = m(\frac{x}{\hat{p}} - \frac{1}{2}\{\hat{q}, \frac{1}{\hat{p}}\}_+) = -e^{-i\hat{p}x}\sqrt{\frac{m}{\hat{p}}} \, \hat{q} \, \sqrt{\frac{m}{\hat{p}}} \, e^{i\hat{p}x}.$$
(9)

Notice the proliferation of carets above. It is a reminder that we now deal with operators acting on the Hilbert space of the free particle states. From now on, we will drop the operator carets, simplifying the notation as much as possible, wherever this will not produce confusion between operators and *c*-number variables. The eigenstates $|txs0\rangle$ of this operator $t_0(x)$ in the momentum representation can be given as $(\hbar = 1)$

$$\langle p|txs0\rangle = \theta(sp)\sqrt{\frac{|p|}{m}} \exp(i\frac{p^2}{2m}t)\langle p|x\rangle,$$
 (10)

where t is the time eigenvalue, x the arrival position, and where we use s = r for right-movers (p > 0), and s = l for left-movers (p < 0.) The label 0 stands for free particle case. Finally, the argument sp of the step function that appears on the rhs is +p for s = r, and -p for s = l, so that

$$\theta(rp) = \int_0^\infty dp |p\rangle \langle p| \text{ and } \theta(lp) = \int_{-\infty}^0 dp |p\rangle \langle p|.$$
(11)

The degeneracy of the energy with respect to the sign of the moment is explicitly shown by means of a label s = r, l in the energy representation, where

$$\langle Es'0|txs0\rangle = \delta_{s's} \left(\frac{2E}{m}\right)^{1/4} e^{iEt} \langle Es0|x\rangle.$$
(12)

Summarizing, there is a representation for the time of arrival at x spanned by the eigenstates

$$|txs0\rangle = (\frac{2H_0}{m})^{1/4} e^{iH_0 t} \Pi_{s0} |x\rangle,$$
(13)

where Π_{s0} projects on the subspace of right-movers (s = r), or of left-movers (s = l), *i.e.*,

$$\Pi_{s0} = \int_0^\infty dE |Es0\rangle \langle Es0| = \theta(sp). \tag{14}$$

These time eigenstates are not orthogonal. This gave rise in the past to serious doubts about their physical meaning. The origin of the problem can be traced back to the fact that (9) is not self-adjoint, that is, that $\langle \varphi | t_0(x) \psi \rangle \neq \langle t_0(x) \varphi | \psi \rangle$. This was

pointed out by Pauli [7] a long time ago and is due to the lower bound on the energy spectrum that prevents the applicability of the Stone theorem [16]. The problem emerges as soon as one attempts integration by parts in the energy representation.

Not being self-adjoint or orthogonal, this operator poses an interpretation problem that can be solved by considering it in terms the Positive Operator Valued Measures (POVM). This is a class of operators less restrictive than the traditional projector valued measures. The POV measures only requires the hermiticity of $t_0(x)$ (*i.e.* $t_0(x) = (t_0(x))^{*\top}$) to assure the positivity of the measure. Now, instead of a Projector Valued spectral decomposition of the identity operator, one has the POV measure

$$P_0(\Pi(x); t_1, t_2) = \sum_s \int_1^2 dt \, |txs0\rangle \, \langle txs0|$$
$$= \sum_s \int_1^2 dt \, (\frac{2H_0}{m})^{1/4} \, e^{iH_0 t} \, \Pi_{s0} \, \Pi(x) \, \Pi_{s0} \, e^{-iH_0 t} \, (\frac{2H_0}{m})^{1/4}, \tag{15}$$

whose notation indicates the arrival interval and that the dependence on the arrival position comes through the projector $\Pi(x) = |x\rangle \langle x|$ on the position eigenstate. For the above measure $P_0(1,2)^2 \neq P_0(1,2)$ because $|txs0\rangle\langle txs0|$ is not a projector, as the states are not orthogonal. However, the limit as $t \to \infty$ of $P_0(-t,+t)$ is the identity as can be checked explicitly. The time operator obtained is well suited for interpretation. This solution was introduced in [14], and extensively analyzed in refs. [17, 18]. It has been recently reviewed in [19] and criticized in [20].

In this formulation the time of arrival is given by the first moment of the measure

$$t_0(H_0, \Pi(x)) = \sum_s \int_{-\infty}^{+\infty} dt \, t \, |txs0\rangle \, \langle txs0|$$

= $\int_{-\infty}^{+\infty} dt \, t \, (\frac{2H_0}{m})^{1/4} \, e^{iH_0 t} \, \mathcal{P}_0(x) \, e^{-iH_0 t} \, (\frac{2H_0}{m})^{1/4},$ (16)

where $\mathcal{P}_0(x) = \sum_s \prod_{s0} \prod(x) \prod_{s0}$, which is not a projector. We now have the tools necessary for the physical interpretation of the formalism: Given an arbitrary state ψ at t = 0, its time of arrival at a position x has to be, according to (16),

$$\langle \psi | t_0(x) | \psi \rangle = \frac{1}{P_0(x)} \sum_s \int_{-\infty}^{+\infty} dt \, t \, |\langle txs0 | \psi \rangle|^2, \tag{17}$$

with the standard interpretation of $\sum_{s} |\langle txs0|\psi\rangle|^2$ like the (as yet unnormalized) probability density that the state $|\psi\rangle$ arrives at x in the time t. The probability of arriving at x at any time is then $P_0(x) = \int dt \sum_{s} |\langle txs0|\psi\rangle|^2$, giving a normalized probability density in times of arrival

$$P_0(t,x) = \frac{1}{P_0(x)} \sum_{s} |\langle txs0|\psi\rangle|^2,$$
(18)

the normalization that has been used in (17).

The above equations (17,18) can be given forms that are very useful for computation and that throw some light on the physical meaning of the different quantities involved. By using explicitly (13), one gets

$$P_{0}(x) = \sum_{s} \{ \int dE \left(\frac{2E}{m}\right)^{1/4} \langle x | Es0 \rangle \langle Es0 | \psi \rangle \}^{*} \\ \times \{ \int dE' \left(\frac{2E'}{m}\right)^{1/4} \langle x | E's0 \rangle \langle E's0 | \psi \rangle \} \int dt \, e^{-i(E-E')t} \\ = 2\pi \sum_{s} \int dE \left(\frac{2E}{m}\right)^{1/2} |\langle x | Es0 \rangle \langle Es0 | \psi \rangle|^{2}.$$
(19)

The use of a similar procedure in (17) leads to

$$\langle \psi | t_0(x) | \psi \rangle = -\frac{i\pi}{P_0(x)} \sum_s \int dE (\frac{2E}{m})^{1/2} \left\{ \langle x | Es0 \rangle \langle Es0 | \psi \rangle \right\}^*$$

$$\times \overleftrightarrow{\frac{\partial}{\partial E}} \left\{ \langle x | Es0 \rangle \langle Es0 | \psi \rangle \right\}$$

$$= \frac{2\pi}{P_0(x)} \sum_s \int dE (\frac{2E}{m})^{1/2} | \langle x | Es0 \rangle \langle Es0 | \psi \rangle |^2$$

$$\times \frac{\partial}{\partial E} \left\{ \arg \langle x | Es0 \rangle + \arg \langle Es0 | \psi \rangle \right\}.$$

$$(20)$$

This expression is easy to understand. In fact it involves two ingredients: the plane wave amplitude $\langle x|Es0\rangle = \sqrt{\frac{m}{2\pi p}} \exp(ispx)$, along with the bracket $\langle Es0|\psi\rangle = \sqrt{\frac{m}{p}}\tilde{\psi}(sp)$ where $\tilde{\psi}$ is the Fourier transform of the initial state in momentum space, and $p = \sqrt{2mE}$. This gives for the arrival amplitude

$$\langle txs0|\psi\rangle = \frac{1}{2\pi} \int_0^\infty dp \,\sqrt{\frac{p}{m}} e^{i(-Et+spx)} \tilde{\psi}(sp). \tag{21}$$

This is a free case, so that the probability of ever arriving at x has to be one. In fact

$$P_0(x) = \sum_{s=r,l} \int_0^\infty dp \ |\tilde{\psi}(sp)|^2 = \int_{-\infty}^{+\infty} dp \ |\tilde{\psi}(p)|^2 = 1,$$
(22)

where we used that in our notation rp = +p and lp = -p. We also have:

$$\frac{\partial}{\partial E} \{ \arg\langle x | Es0 \rangle + \arg\langle Es0 | \psi \rangle \} = \frac{m}{p} (sx - \frac{\partial \arg(\tilde{\psi}(sp))}{\partial p}).$$
(23)

There are initial wave packets centered around the values (q_0, p_0) for which $\tilde{\psi}(p) = |\tilde{\psi}(p)| \exp(-ipq_0)$ with the amplitude $\tilde{\psi}(p)$ peaked around p_0 . Then,

$$(\partial/\partial p) \arg \tilde{\psi}(sp) = sq_0,$$

and the time of arrival at x reduces to

$$\langle \psi | t_0(x) | \psi \rangle = \sum_{s=r,l} \int_0^\infty dp \, |\tilde{\psi}(sp)|^2 \, \frac{m(x-q_0)}{sp}$$

= $\int_{-\infty}^{+\infty} dp \, |\tilde{\psi}(p)|^2 \, \frac{m(x-q_0)}{p},$ (24)

which is the time of arrival of the classical free particle averaged over its initial state.

The generalization to the case of three space dimensions [13] is not straightforward. The reason is that to begin with, there are three equivalent equations (7) for the time of arrival. To be compatible, they have to satisfy the constraints

$$\mathcal{L} = (\mathbf{q} - \mathbf{x}) \wedge \mathbf{p} = 0, \tag{25}$$

where we drop the distinctions made in (7) between upper and lower case letters, as they are the same objects for free particles. Classically, the constraints correspond to the fact that **x** has to be a point of the particle's trajectory, therefore the angular momentum can be written as $\mathbf{L} = \mathbf{x} \wedge \mathbf{p}$. In other words, the angular momentum with respect to **x**, that is \mathcal{L} , has to vanish. We now show that the constraints (25) are first class. First of all, they are closed as their components $\mathcal{L}_a = \epsilon_{abc} (q - x)_b p_c$ satisfy the algebra of 3-D rotations, namely $\{\mathcal{L}_a, \mathcal{L}_b\} = \epsilon_{abc} \mathcal{L}_c$. Then, the total Hamiltonian is $H_T = \frac{\mathbf{p}^2}{2m} + \lambda \cdot \mathcal{L}$, where λ is a vector multiplier, so that $\{\mathcal{L}_a, H_T\} =$ $\epsilon_{abc} \lambda_b \mathcal{L}_c$. Therefore, the constraints form a first class system that depends parametrically on **x**, one for each arrival position **x**. Not all the **x**'s can be reached from an arbitrary set (**q**, **p**) of phase space variables. Only those **x** that satisfy the constraints are positions where the particles with these dynamical variables can eventually be detected. A detector placed somewhere else will miss them.

The above translates into quantum mechanics as it is: Not all the states in the Hilbert space of free particle states \mathcal{H} with Hamiltonian H_0 can be detected at a specific position \mathbf{x} . Only the subspace $\mathcal{H}_{\mathbf{x}}$ of the states that satisfy the constraints (25) (where now \mathbf{q} and \mathbf{p} are operators) qualify as the Hilbert space of detected states (at \mathbf{x}). This subspace is spanned by the states $|\psi; \mathbf{x}\rangle \in \mathcal{H}$ of the form $|\psi; \mathbf{x}\rangle = \psi(H_0) |\mathbf{x}\rangle$. Here, $\psi(H_0)$ is an arbitrary function of H_0 , that may also depend on other parameters, $|\mathbf{x}\rangle$ is the eigenstate $\mathbf{q} |\mathbf{x}\rangle = \mathbf{x} |\mathbf{x}\rangle$ of the arrival position. In particular, the detected subspace $\mathcal{H}_{\mathbf{x}}$ is obtained from \mathcal{H}_0 by a translation of amount \mathbf{x} , as required by covariance.

The value of t comes from the equation of motion in the subspace orthogonal to the constraints, namely $\mathbf{p} \cdot \mathbf{x} = \mathbf{p} \cdot \left(\frac{\mathbf{p}}{m} t_0(\mathbf{x}) + \mathbf{q}\right)$, that in spherical coordinates where $|\mathbf{p}\rangle = |p, \theta_p, \phi_p\rangle$ and $q = i\frac{d}{dp}$ can be written as $x = \frac{p}{m} t_0(\mathbf{x}) + q$. This can be readily inverted to give

$$t_0(\mathbf{x}) = -\frac{m}{p} e^{-i\mathbf{p}\mathbf{x}} p^{-1/2} q p^{1/2} e^{i\mathbf{p}\mathbf{x}}.$$
 (26)

Notice the characteristic powers of p to the right and to the left of q. This operator ordering makes of t_0 a maximally symmetric operator with respect to the measure d^3p , making integration by parts a straightforward task. In d space dimensions we would have $t_0 \propto \frac{1}{p^{n+1}} \left(-i\frac{d}{dp}\right) p^n$ with n = (d-2)/2 [13]. The eigenfunctions of t_0 are given in the momentum representation by:

$$\langle \mathbf{p}|t; \, \mathbf{x}, \, 0 \rangle = \sqrt{\frac{1}{4\pi m p}} \, e^{iE_{\mathbf{p}}t} \langle \mathbf{p}| \, \mathbf{x} \rangle, \tag{27}$$

where $t \in \mathbf{R}$ is the time eigenvalue, and $E_{\mathbf{p}} = p^2/2m$. One can define a time of arrival representation given by

$$|t; \mathbf{x}, 0\rangle = \frac{1}{\sqrt{4\pi m}} \left(\frac{1}{2 m H_0}\right)^{1/4} e^{iH_0 t} |\mathbf{x}\rangle.$$
(28)

These eigenstates are not orthogonal. They correspond to a POV measure defined by the spectral decomposition

$$\mathbf{1}_{\mathbf{x}} = \int_{-\infty}^{+\infty} dt \, |t; \, \mathbf{x}, \, 0\rangle \langle t; \, \mathbf{x}, \, 0 \, |.$$
⁽²⁹⁾

It can be immediately seen that for any state $|\psi; \mathbf{x}\rangle \in \mathcal{H}_{\mathbf{x}}$, and for arbitrary momentum \mathbf{p}

$$\langle \mathbf{p} \, | \, \mathbf{1}_{\mathbf{x}} \, | \, \psi; \mathbf{x} \rangle = \langle \mathbf{p} \, | \, \psi; \mathbf{x} \rangle \quad \forall \, \mathbf{p} \in \mathbf{R}^3.$$
(30)

Therefore, the operator $\mathbf{1}_{\mathbf{x}}$ is a decomposition of the identity within the subspace of detected states $\mathcal{H}_{\mathbf{x}}$. The fact that $\mathbf{1}_{\mathbf{x}} < \mathbf{1}$, so that the decomposition is uncompleted, is the quantum version of the classical case where only a part of the incoming particles will (reach and) be detected at \mathbf{x} . From here it is clear that our formalism is finer than that provided by the so-called screen operators [22], that would describe the arrival at a two dimensional plane put across the particle trajectories. In fact, these screen operators would correspond to a coarse graining of the present formalism, whose interpretation is analyzed in some detail in [13].

The time of arrival can be given through the first momentum of the POV measure (29):

$$t_{0}(\mathbf{x}) = \int dt \ t \ |t; \ \mathbf{x}, 0\rangle \langle t; \ \mathbf{x}, 0| = \frac{1}{4\pi m} \int dt \ t \ (\frac{1}{2 \, m \, H_{0}})^{1/4} e^{iH_{0} \ t} \, | \ \mathbf{x}\rangle \langle \ \mathbf{x} \, | \, (\frac{1}{2 \, m \, H_{0}})^{1/4} e^{-iH_{0} \ t}, \tag{31}$$

whose similarity with the 1-D case (16) is evident, and can be used as a guide to get the average time of arrival an other quantities of interest that were worked on in one space dimension.

4. The arrival of interacting particles

In this section we want to determine the effect on the times of arrival of a position dependent interaction between the particle and the medium, that we describe by a potential energy V(q). For instance, we want to consider the case of a barrier placed between the detector and the initial state. We would put a detector at x (at the other side of the barrier), and prepare the initial state $|\psi\rangle$ of the particle at t = 0 (at this side of the barrier). We would then record with a clock the time t when the detector clicks. Repeating this procedure with identically prepared initial states, we would get the probability distribution P(t, x) in times of arrival at x. This is the same procedure used for the free particle case, the differences coming from the presence of the potential energy V(q).

To find the quantum time of arrival we will use what we know from the classical case: There is a canonical transformation from the free $(H_0 = \frac{P^2}{2m})$ translation variables (Q, P) to the actual variables (q, p) of the interacting situation where $H = \frac{p^2}{2m} + V(q)$. Time can be given equally in any of these two versions, and we did already quantize the free version t_0 in the previous section. Now, in successive steps, we do the following [21]: We first construct the quantum canonical transformation U that connects the free-particle states to the eigenstates of the complete Hamiltonian. This is the quantum version of the (inverse of the) Jacobi-Lie canonical transformation (2,5). We will see later on that U is given by the Möller wave operator. We will then apply U to t_0 to define the time of arrival t in the presence of the interaction potential V(q) in terms of t_0 . We will also address some questions of interpretation of the resulting formalism.

Dirac introduced canonical transformations in quantum mechanics in a number of different places [23] by means of unitary transformations $U(UU^{\dagger} = U^{\dagger}U = 1)$. To fix the notation, we assume in what follows that the operators q and p are given in the coordinate representation of the Hilbert space $L^2(x)$ by q = x and $p = -i\hbar \frac{\partial}{\partial x}$. If the operators \bar{q} and \bar{p} are the result of an arbitrary canonical transformation applied to q and p, then there is a unitary transformation U such that

$$\bar{q} = U^{\dagger} q U, \ \bar{p} = U^{\dagger} p U \Rightarrow \ [\bar{q}, \bar{p}] = [q, p] = -i\hbar.$$
(32)

One can also define implicitly the quantum canonical transformations as is done in classical mechanics, a possibility that has been thoroughly analyzed and developed. The main results of the method are collected in [24], which also includes references to other relevant literature. The definition of U is given implicitly by the two conditions

$$F(\bar{q}, \bar{p}) = F_0(q, p), \text{ and } G(\bar{q}, \bar{p}) = G_0(q, p),$$
 (33)

where F, G, F_0 and G_0 are functions of the operators shown explicitly as their arguments. They cannot be chosen arbitrarily, the necessary and sufficient condition for

the canonicity of the transformation being $[F, G] = [F_0, G_0]$. The dependence of (33) on U can be explicitly given by using (32) in it:

$$U^{\dagger} F(q,p) U = F_0(q,p), \text{ and } U^{\dagger} G(q,p) U = G_0(q,p),$$
 (34)

that comes from the straight application of (32) to the first members. In addition, U is unitary so that the spectra of the original and transformed operators have to coincide. We now assume that F and F_0 are self-adjoint operators whose eigenvalue problems are solved by the states $|f s\rangle$ and $|f s 0\rangle$ (both corresponding to the same eigenvalue), that form orthogonal and complete bases of the Hilbert space satisfying

$$F | f s \rangle = \lambda_f | f s \rangle, \quad F_0 | f s 0 \rangle = \lambda_f | f s 0 \rangle. \tag{35}$$

We are accepting here the presence of degeneracy indicated by the discrete index s, something that we will need later. Assuming now a continuous spectrum (the case we will be interested in), the operator U that satisfies the first row of (35) is given by

$$U = \sum_{s} \int_{\sigma(\lambda)} d\lambda_f |fs\rangle \langle fs0|.$$
(36)

It is straightforward to verify that it is unitary. We can now give the definition of G in terms of G_0 using U, that is $G = U G_0 U^{\dagger}$, which in full detail reads

$$G(q,p) = \sum_{ss'} \int_{\sigma(\lambda)} d\lambda_f d\lambda_{f'} |fs\rangle \langle fs0|G_0(q,p)|f's'0\rangle \langle f's'|.$$
(37)

This is the main result of our procedure. The fact that we can define an operator G, canonically conjugate to F, if we know G_0 and U.

We will now apply this to the case where F_0 is the free Hamiltonian H_0 , F the complete Hamiltonian H and G_0 the time of arrival t_0 of the free particle Eqs.(9), or (16). Then, we have $H_0 = U^{\dagger} H U$ and $\Pi_0(x) = U^{\dagger} \Pi(x) U$. Associated to the free particle there was the positive operator valued measure P_0 of Eq.(15). Accordingly, the POV measure P of the interacting case will be given by (cf(34))

$$P(\Pi(x); t_1, t_2) = U P_0(\Pi_0(x); t_1, t_2) U^{\dagger},$$
(38)

and the time of arrival operator in the presence of interactions (the G of (37)) is given by

$$t(H, \Pi(x)) = Ut_0(H_0, \Pi_0(x))U^{\dagger}.$$
(39)

We noticed above that the spectra of the original and transformed operators had to coincide. Now, $\sigma(H_0) = \mathbf{R}^+$ so that not all the Hamiltonians can be obtained from H_0 by this procedure. In general, some fixing will be required to make the spectra coincide. Here we will only consider well behaved potentials ($V(q) \ge 0 \forall q \in \mathbf{R}$), vanishing appropriately at spatial infinity. This ensures the required coincidence of

the spectra, but introduces two solutions for U due to the existence of two independent sets of eigenstates of H:

$$U_{(\pm)} = \sum_{s} \int_{0}^{\infty} dE |Es(\pm)\rangle \langle Es0| = \Omega_{(\pm)}.$$
(40)

These are the Möller operators connecting the free particle states to the bound and scattering states. In the presence of bound states these operators would only be isometric, because the correspondence between eigenstates of H and free states could not be one-to-one. In our case $V(q) \ge 0$, there is one free state for each scattering state and conversely. Thence, the Möller operators are unitary. In this case, the intertwining relations $H\Omega_{(\pm)} = \Omega_{(\pm)}H_0$ can be put in the more desirable form $H = \Omega_{(\pm)}H_0\Omega_{(\pm)}^{\dagger}$. We will also follow the standard sign conventions, choosing $\Omega_{(+)}$ in (40) that, when $E = \lim_{\epsilon \to 0^+} (E + i\epsilon)$, gives signal propagation forward in time. The results that would be obtained with $\Omega_{(-)}$ would correspond to the time reversal of this situation. If τ is the time reversal operator, then $P_{(-)}(\Pi(x); t_1, t_2) = \tau P_{(+)}(\Pi(x); -t_2, -t_1) \tau^{\dagger}$. For notation simplicity, we will omit these labels (\pm) wherever possible.

The parameter x that appears in (37) and (39) is the actual detection position in the interacting case, the place whose time of arrival at we want to know. Therefore, the arguments of t in (39) have to be $\Pi(x) = |x\rangle\langle x|$ and H. Hence, the argument of t_0 will be an object $\Pi_0(x) = \Omega^{\dagger}\Pi(x)\Omega$ which collects all the states of the free particle that add up to produce the actual position eigenstate $|x\rangle$ by the canonical transformation. Much of the difference between the classic and quantum cases is hidden here. In particular, the quantum capability to undergo classically forbidden jumps in phase space has much to do with the fact that $\Pi(x)$ and $\Pi_0(x)$ cannot be position projection operators simultaneously.

We have now at hand all the tools necessary to answer the questions about the time of arrival of interacting particles. Given a particle that was initially (at t = 0) prepared in the state $|\psi\rangle$, we can compute the predictions for the average time of arrival $\langle \psi | t(x) | \psi \rangle$, the probability distribution in times of arrival P(t, x) and the probability of ever arriving at x, P(x). Instead of writing more equations, we refer the reader to Eqs. (16,17), (15,18) and (19). By simply erasing the label 0 from them, one gets the correct expressions for the interacting case, with the caveat that – to be of practical use – they require the knowledge of the scattering states and Möller operator. It is worth recalling here that the expression (20) for the average time remains valid after dropping the 0's. So, $\langle \psi | t(x) | \psi \rangle$ is still the sum of two independent pieces, one containing $(\partial/\partial E) \arg \langle x | E s \rangle$ and only depends of the position of arrival.

We now consider the case where there is a finite potential energy starting at the origin $(V(q) = 0, \forall q \leq 0)$, which is so smooth that the quasi-classical approxima-

tion is valid. Then for E > V(x) the exponentially small reflection amplitude can be neglected, giving the scattering states

$$\langle x|E\,r\,\rangle \approx \theta(-x)\,\sqrt{\frac{m}{2\pi p}}\,e^{ipx}\,+\,\theta(x)\,\sqrt{\frac{m}{2\pi p(x)}}\,e^{i\int_{0}^{x}dq\,p(q)},$$
 (41)

with $p(q) = \sqrt{2m(E - V(q))}$, that are normalized to an incoming right-moving particle by time unit. We now consider the physically interesting case where the initial wave packet is normalized to 1, (*i.e.* that $\int dp |\tilde{\psi}(p)|^2 = 1$ with $\tilde{\psi}(p) = \langle p | \psi \rangle$), also, that it is localized around a position q_0 well to the left of the origin, and that it has a mean momentum $p_0 \gg V(x)$. Then, to this order the probability of ever arriving at x (*c.f.* (19)) gives

$$P(x) \approx \theta(-x) + \theta(x) P_{>}(x), \text{ where } P_{>}(x) = \int_{0}^{\infty} dp \frac{p}{p(x)} |\tilde{\psi}(p)|^{2} \qquad (42)$$

so that $\frac{p}{p(x)} |\tilde{\psi}(p)|^2$ is the (unnormalized) probability of arrival at the point x with momentum p(x), as corresponds to the quasiclassical case. Notice that to the left of the origin the result is the same as in the free case. This comes about because the approximation neglects reflection, thus missing at q < 0 any information about the existence of a finite V at $q \ge 0$. For the time probability distribution one gets

$$P(t,x) \approx \frac{\theta(-x)}{2\pi} \left| \int_0^\infty dp \, e^{-iEt} \, \tilde{\psi}(p) \right|^2 + \frac{\theta(x)}{2\pi P_>(x)} \left| \int_0^\infty dE \, \sqrt{\frac{m}{p(x)}} \, \tilde{\psi}(p) \, e^{-i(Et - \int_0^x dq \, p(q))} \right|^2,$$
(43)

which, not surprisingly, is the same as that of free particles for x < 0. Finally,

$$\langle \psi | \mathbf{t}_x | \psi \rangle \approx \theta(-x) \int_0^\infty dp \, |\tilde{\psi}(p)|^2 \, \frac{m}{p} \{ x - q_0 \}$$

$$+ \frac{\theta(x)}{P_>(x)} \int_0^\infty dp \, \frac{p}{p(x)} \, |\tilde{\psi}(p)|^2 \, \{ -\frac{mq_0}{p} + m \int_0^x \frac{dq}{p(q)} \}.$$

$$(44)$$

Therefore, we recover the time of arrival of the free particles for negative x. On the other hand, for x > 0 we get the classical time of arrival at x for initial conditions $(q_0, p), \int_{a_0}^{x} (m/p(q)) dq$, weighted by the probability of these conditions.

5. Advanced or delayed arrival?

What is the effect of putting a quantum barrier in the path of the arriving particle? Hartman [25] studied this question a long time ago, reaching the conclusion that tunneled particles should appear instantaneously on the other side of the barrier. Our formalism supports this result, but only for thin enough barriers.

The time of arrival at a point x in the presence of a barrier will be given through a probability amplitude

$$\langle txs|\psi\rangle = \int dE \left(\frac{2E}{m}\right)^{1/4} e^{-iEt} \langle x|Es(+)\rangle \langle Es(+)|\psi\rangle. \tag{45}$$

In the case where x is at the right of the barrier, the amplitude can be approximately given by [21]

$$\langle txs|\psi\rangle \approx \frac{\delta_{sr}}{\sqrt{2\pi}} \int dE (\frac{m}{2E})^{1/4} e^{-i(Et-px)} T(p) \tilde{\psi}(p),$$
 (46)

where T(p) is the transmission amplitude for momentum p. Now, the total probability of eventually arriving at x in any time t is $P(x) \approx \int dp |T(p) \tilde{\psi}(p)|^2$ that is independent of x in cases like this, where x is beyond the range of the potential. After a straightforward calculation we get for the average time of arrival at the other side of the barrier the corresponding version of (20)

$$\langle \psi | \mathbf{t}_x | \psi \rangle \approx \frac{1}{P(x)} \int_0^\infty dp \, |T(p) \, \tilde{\psi}(p)|^2 \frac{m}{p} \{ x - q_0 + \frac{d \arg(T(p))}{dp} \}. \tag{47}$$

It is the value of the Wigner time [26] averaged over the transmitted state.

Consider a simple square barrier of height V and width a. The transmission coefficient is in this case:

$$T(p) = e^{-ipa} \left(1 - i \frac{(p^2 + p'^2)}{2pp'} \tan p'a \right)^{-1} \sec p'a,$$
(48)

where $p' = \sqrt{p^2 - p_V^2}$, that is imaginary for p below p_V . Notice the contribution -pa to $\arg(T(p))$. This will subtract a term a to the path length $x - q_0$ that appears in (47). The barrier has effective zero width or, in other words, it is traversed instantaneously. This is the Hartman effect for barriers. To be precise, the effect is not complete, it is compensated by the other dependences in p'a present in the phase of T(p). In fact, it disappears for low barriers $(p_V/p) \rightarrow 0$, where all the a dependences of the phase cancel out, as was to be expected because the barrier effectively vanishes in this limit. In the opposite case of high barriers $(p/p_V) \rightarrow 0$ the effect saturates and there is a decrease $-\frac{ma}{p}$ in the time of arrival of transmitted plane waves, which emerge almost instantaneously at the other side of the barrier.

The averaging of the Wigner time over the transmitted state, present in (47) as a consequence of the formalism, has dramatic effects, because it effectively forbids the transmission of the wave components with low momenta. In fact, it produces the exponential suppression (by a factor $\exp(-2|p'|a)$) of the tunneled components. Therefore, only the components with momentum above p_V have a chance of surmounting thick barriers, being finally transmitted. But these components are delayed by the barrier (for them $(\frac{d \arg(T(p))}{dp}) > 0$), whose overall effect transmutes from advancement into retardation [21] at a definite predictable thickness that depends on the barrier height and also on the properties of the incoming state.

Acknowledgements

Parts of this work are based on a talk given at the I Cracow-Clausthal Workshop on Tunneling Effect and Other Fundamental Problems in Quantum Physics. The author is indebted to Jacek Jakiel and Edward Kapuścik for the warm hospitality extended to him during the Workshop. This work has been partially supported by the DGESIC under contract PB97-1256.

References

- V. I. Arnold, K. Vogunann and A. Weinstein, Mathematical Methods of Classical Mechanics, Springer Verlag, Berlin, 1989.
- [2] J. L. McCauley, Classical Mechanics, Cambridge University Press, Cambridge, 1998.
- [3] A. Einstein, Verh. Dtsch. Phys. Ges. 19 (1917) 82.
- [4] P. S. Epstein, Verh. Dtsch. Phys. Ges. 19 (1917) 116.
- [5] M. C. Gutzwiller, Chaos in Classical and Quantum Mechanics, Springer Verlag, Berlin, 1990.
- [6] E. T. Whittaker, A treatise on the analytical dynamics of particles and rigid bodies, Cambridge University Press, Cambridge 1956.
- [7] W. Pauli, in Encyclopedia of Physics Vol. 5, ed. S. Flugge, Springer, Berlin 1958.
- [8] Y. Aharonov and D. Bohm, Phys. Rev. 122 (1961) 1649.
- [9] H. Paul, Ann. Phys. (Leipzig) 7 (1962) 252.
- [10] J. Kijowski, Rep. Math. Phys. 6 (1964) 361.
- [11] E. A. Galapon, quant-ph/9908033 and quant-ph/0001062.
- [12] N. Grot, C. Rovelli and R. S. Tate, Phys. Rev. A 54 (1996) 4676.
- [13] J. León, J. Phys. A 30 (1997) 4791, quant-ph/9608013.
- [14] R. Giannitrapani, Int. J. Theor. Phys. 36 (1997) 1575.
- [15] V. Delgado and J. G. Muga, Phys. Rev. A 56 (1997) 3425.
- [16] M. H. Stone Linear transformations in Hilbert space Am. Math. Soc. New York 1932, 1979.
- [17] R. Giannitrapani, J. Math. Phys. 39 (1998) 5180.
- [18] M. Toller, Phys. Rev. A 59 (1999) 960, quant-ph/9702060.
- [19] J. G. Muga, C. R. Leavens, and J. P. Palao, Phys. Rev. A 58 (1998) 4336, quant-ph/9807066.
- [20] J. Kijowski, Phys. Rev. A 59 (1999) 897.
- [21] J. León, J. Julve, P. Pitanga and F. J. de Urríes, Phys. Rev. A 61 (2000) 062101, quant-ph/0002011.
- [22] B. Mielnik, Found. Phys. 24 (1994) 1113.
- [23] P. A. M. Dirac, *The Principles of Quantum Mechanics*, 4th Ed., Clarendon Press, Oxford 1959, and *Phys. Z. Sowj.* Band 3, Heft 1 (1933), reprinted in "Selected papers on Quantum Electrodynamics", Ed. J. Schwinger, Dover, New York, 1958.
- [24] M. Moshinsky and T. H. Seligman, Ann. Phys. NY 114 (1978) 243; ibid 120 (1979) 402.
- [25] T. E. Hartman, J. Appl. Phys. 33 (1962) 3427.
- [26] E. P. Wigner, Phys. Rev. 98 (1955) 145.

Superluminal Phenomena and the Phenomenological Maxwell Equations

Jakub Rembieliński Department of Theoretical Physics, University of Łódź, ul. Pomorska 149/153, 90-236 Łódź, Poland

Motivated by a number of recent experiments [1, 2, 3, 4, 5], in this paper we discuss solutions of effective Maxwell-like equations describing the propagation of an electromagnetic field in a medium that "feels" a quantum preferred frame.

Keywords: preferred frame, superluminal, absolute synchronization

1. Introduction

As is well known, from the "orthodox" point of view there is a "peaceful coexistence" between SR and QM if a physical meaning is attributed to final probabilities only [6, 7, 8]. However, such a restrictive approach is unsatisfactory for many physicists, for whom also the notion of a physical state, its time evolution, localization of quantum events, *etc.* should have a "real" and not just a technical meaning.

According to this second approach to understanding QM we encounter a number of theoretical problems on the borderline between QM and SR. The most important ones are related to the apparent nonlocality of QM and lack of a manifest Lorentz covariance of quantum mechanics of systems with finite degrees of freedom. Recently, several authors have suggested that a proper formulation of QM needs the introduction of a preferred frame (PF) [9, 10, 11, 12]. In particular, introduction of a PF can solve some dilemmas relating to the causal description of quantum collapse in the EPR-like experiments with moving reference frames [13]. It is important to stress that the notion of a PF used here is completely different from the traditional notion, linked to the ether, and is in agreement with classical experiments. Most recent EPR experiments performed in Geneva [14] do not contradict the PF hypothesis and give a lower bound for the speed of "quantum information" in the cosmic background radiation frame (CBRF) at 1.5×10^4c .

A conceptual difficulty related to the PF notion lies in an apparent contradiction with the Lorentz symmetry. But as was shown in the [12, 15, 16] this is not the case: it is possible to arrange the Lorentz group transformation in such a way that the Lorentz covariance survives while the relativity principle (democracy between inertial frames) is broken. Moreover, such approach is consistent with the classical phenomena. Recall also that attention was recently devoted to the PF as a consequence of a possible breaking of the Lorentz invariance [17, 18] in high energy physics. We are not so "radical" in this paper because it is enough to break the relativity principle only in order to extend the causality notion and consequently to reconcile QM with the Lorentz covariance.

We introduce and discuss a direct generalization of the macroscopic (phenomenological) Maxwell equations which are both Lorentz-covariant and "feel" the preferred frame. We show that, according to these equations, the electromagnetic field propagates faster than light in vacuum, *i.e.*, the effective mass of the photon is tachyonic. Although our derivation is purely classical, it is motivated by the fact that in a medium, light propagation is mainly a quantum phenomenon; therefore the influence of the PF (if it really exists) can in principle be observed. In the following we make simplifying assumptions, such as homogeneity and isotropy of the medium.

Because a "folk theorem" which states that local Lorentz covariance implies relativity (*i.e.*, democracy between inertial frames) is commonly used, we begin with a brief review of the formalism introduced in [12, 15, 16]. Obviously, if we try to realize the Lorentz group as a linear transformation of the Minkowski coordinates only, the above mentioned "theorem" is necessarily true. However, if a PF is distinguished, we have at our disposal an additional set of parameters, namely the four-velocity of the PF with respect to each inertial observer. Using this freedom we can realize the Lorentz group in a nonstandard way [15, 16]. Physically, such a realization of the Lorentz transformations corresponds to a nonstandard choice of the synchronization scheme for clocks [19]. In [12] this scheme was applied to formulating the manifestly covariant QM.

To be concrete, in that approach the Lorentz group is realized in a standard way insofar as it is restricted to rotations, while for boosts we have

$$\begin{aligned} x^{\prime 0} &= \frac{1}{w^0} x^0, \\ \mathbf{x}^{\prime} &= \mathbf{x} - \mathbf{w} \left(x^0 + u^0 (\mathbf{u} \cdot \mathbf{x}) - \frac{\mathbf{w} \cdot \mathbf{x}}{1 + \sqrt{1 + \mathbf{w}^2}} \right), \end{aligned}$$
(1)

and

$$u'^{0} = \frac{1}{w^{0}}u^{0},$$

$$\mathbf{u}' = \mathbf{u} - \mathbf{w} \left(\frac{1}{u^{0}} - \frac{\mathbf{u} \cdot \mathbf{w}}{1 + \sqrt{1 + \mathbf{w}^{2}}}\right),$$
(2)

where $u^{\mu} = (u^0, \mathbf{u})$ and $w^{\mu} = (w^0, \mathbf{w})$ is the (timelike) four-velocity of the PF and $[x'^{\mu}]$, respectively as observed from the inertial frame $[x^{\mu}]$. The four-vectors u^{μ} and

 w^{μ} are related to three-velocities via the following formulae

$$\mathbf{v} = \frac{\mathbf{u}}{u^0},$$

$$\mathbf{V} = \frac{\mathbf{w}}{w^0},$$

$$\frac{1}{v^0} = \sqrt{(1+u^0\mathbf{u}\cdot\mathbf{V})^2 - \mathbf{V}^2},$$
(3)

The explicit relationship to the standard (Einstein-Poincaré) synchronization is given by $x_E^0 = x^0 + u^0 \mathbf{u} \cdot \mathbf{x}$, $\mathbf{x}_E = \mathbf{x}$, so the time lapse at a space point is the same in both synchronizations. Furthermore, the average light speed over closed loops is constant and equal to the speed of light in vacuum (here c = 1) in agreement with Michelson-like experiments. It is important to stress that both synchronizations (Einstein Poincaré and the nonstandard one) lead to the same results for velocities less than or equal to the speed of light, but only the nonstandard synchronization scheme can be used for a consistent description of possible superluminal phenomena [16]. This is because (as we see from (1)–(2) in the nonstandard synchronization the Lorentz transformations have a triangular form, so the zeroth component of a covariant four-vector is rescaled by a positive factor only. Consequently, an absolute notion of causality can be introduced in this framework. Moreover, if superluminal propagation of information does exist in reality, a PF *must be* distinguished, and consequently, a convention of synchronization as well as the relativity principle are broken. An exhaustive discussion of the nonstandard formulation of Lorentz covariance in this language is given in [12, 16].

2. Effective Maxwell equations

In a homogeneous and isotropic medium the fields **D** and **H** are related to **E** and **B** via permittivity ε^{-1} and permeability μ , respectively, where ε and μ are nonlinear functionals of $\mathbf{E}^2 - \mathbf{B}^2$ and $\mathbf{E} \cdot \mathbf{B}$, in a nonlocal way. To simplify our considerations as far as possible, let us assume that in some range of field intensity ε and μ vary slowly, so they can be treated approximately as constants. Therefore, in our equations we will use only **E** and **B**, *i.e.*, the electromagnetic field tensor $F^{\mu\nu}$ and its dual $\hat{F}^{\mu\nu} = \frac{1}{2}\varepsilon^{\mu\nu\sigma\lambda}F_{\sigma\lambda}$. Moreover, we assume that the possible (quantum) response of the medium, related by preference by QM of a PF, is roughly speaking proportional to **E** and **B**. Under such extremelly simplified assumptions our phenomenological Maxwell-like equations take the form

$$\partial_{\mu}F^{\mu\nu} + \alpha u_{\mu}F^{\mu\nu} = j^{\nu}, \qquad (4)$$

$$\partial_{\mu}\hat{F}^{\mu\nu} + \beta u_{\mu}\hat{F}^{\mu\nu} = 0, \qquad (5)$$

where α and β are constants. In the following we will omit the induced current j^{ν} to concentrate on the consequences of the influence of the PF only. It is not difficult to

check that equations (4–5) with $j^{\nu} = 0$ have nontrivial solutions, admitting a Fourier expansion, only for $\beta = -\alpha$, so (4–5) must be replaced by¹

$$\partial_{\mu}F^{\mu\nu} + \alpha u_{\mu}F^{\mu\nu} = 0, \qquad (6)$$

$$\partial_{\mu}\hat{F}^{\mu\nu} - \alpha u_{\mu}\hat{F}^{\mu\nu} = 0, \qquad (7)$$

with α depending on the properties and the state of the medium. Of course, we can choose $\alpha \ge 0$. Notice that (7) cannot be transformed to the form $\partial_{\mu} \hat{F}^{\mu\nu} = 0$ by a duality transformation. Obviously (6 – 7) are covariant under transformations (1)–(2). Furthermore (6 – 7) necessarily leads to the *tachyonic* wave equation

$$\Box F^{\mu\nu} = \alpha^2 F^{\mu\nu},\tag{8}$$

where $\Box = g^{\mu\nu}(u)\partial_{\mu}\partial_{\nu}$. In the vacuum $\alpha_{\rm vac} = 0$ (more precisely $\alpha_{\rm vac} < 2 \times 10^{-16}$ eV [20]). However in a "PF feeling" medium α should be different from zero.

As was shown in [16] Eq. (8) can be consistently quantized in the nonstandard synchronization, and the resulting theory is not plagued by pathologies relating to quantization of tachyonic field in the SR. In particular, in this framework the vacuum is stable [16]. It is related to the fact that the spectral condition $k^0 > 0$ is invariant also for the space-like dispersion relation $k^2 < 0$ (see transformation law (1)). A covariant construction of the Fock space can also be made [16].

It is easy to see that, using (7) $F^{\mu\nu}$ can be expressed by four-potential A^{μ} as

$$F^{\mu\nu} = \partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} - \alpha(u^{\mu}A^{\nu} - u^{\nu}A^{\mu}), \qquad (9)$$

and the gauge transformations of A^{μ} are of the form $A^{\mu} \rightarrow A^{\mu} + (\partial^{\mu} - \alpha u^{\mu})\chi$. Therefore, the above field equations can be derived from the Lagrangian density

$$L = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}F_{\mu\nu}\left[\partial^{\mu}A^{\nu} - \partial^{\nu}A^{\mu} - \alpha(u^{\mu}A^{\nu} - u^{\nu}A^{\mu})\right].$$
 (10)

For a general field $F^{\mu\nu}$ and under standard identification of $F^{\mu\nu}$ with **E** and **B** $(F^{0k} = E^k, F^{ij} = \varepsilon^{ijk}B^k)$ the Lorentz invariants $F\hat{F}$ and F^2 are

$$\begin{aligned}
F^{\mu\nu}\hat{F}_{\mu\nu} &= -\mathbf{4E}\cdot\mathbf{B}, \\
F^{\mu\nu}F_{\mu\nu} &= -\mathrm{Tr}(gFgF) = \\
&= 2(\mathbf{B}^2 - \mathbf{E}^2) + 4u^0\mathbf{u}\cdot(\mathbf{B}\times\mathbf{E}) - 2(u^0)^2(\mathbf{u}\times\mathbf{B})^2.
\end{aligned}$$
(11)

Now let us examine the monochromatic plane wave solutions $f^{\mu\nu}$ of (6) and (7). Let

$$f^{\mu\nu} = e^{\mu\nu}(k)e^{ikx} + e^{*\mu\nu}(k)e^{-ikx},$$
(12)

¹This can be done equivalently by an appropriate redefinition $F^{\mu\nu} \to \exp(\gamma ux)F^{\mu\nu}$.

where $kx = k_{\mu}x^{\mu}$. Therefore, by (6)and (7) we find

$$\begin{array}{rcl} (ik_{\mu} + \alpha u_{\mu})e^{\mu\nu} &=& 0, \\ (ik_{\mu} - \alpha u_{\mu})\hat{e}^{\mu\nu} &=& 0, \end{array} \tag{13}$$

and (13) lead to the tachyonic dispersion relation $k^2 = -\alpha^2$. The solution of the system (13) has the form

$$e^{\mu\nu} = \left(\frac{\alpha(un) + i(kn)}{\alpha + i(uk)}\right) \times (k^{\mu}u^{\nu} - k^{\nu}u^{\mu}) - (k^{\mu}n^{\nu} - k^{\nu}n^{\mu}) - i\alpha(u^{\mu}n^{\nu} - u^{\nu}n^{\mu}),$$
(14)

where $k^{\mu}, u^{\mu}, n^{\mu}$ and $\varepsilon^{\mu\nu\sigma\lambda}k_{\nu}u_{\sigma}n_{\lambda}$ span a basis, $un = u_{\mu}n^{\mu}$, etc. and n^{μ} can be complex in general.

It is convenient to consider our plane wave solution in the preferred frame. If the PF is realized as the cosmic background radiation frame, this choice is reasonable from our point of view because $v_{\text{solar}} \approx 369.3 \pm 2.5 \text{ km/s} \ll c$ with respect to CBRF [21]. For PF, $u^{\mu} = (1, 0)$ so in this frame $g_{\mu\nu} = \text{diag}(1, -1, -1, -1)$. Now we can put $\mathbf{n} = -(\mathbf{a} + i\mathbf{b})e^{i\varphi}/2$, where \mathbf{a} and \mathbf{b} are real and $\mathbf{a} \perp \mathbf{b}$. Thus from (14) we have the following form of the electromagnetic wave in the preferred frame

$$\mathbf{E} = \frac{1}{|\mathbf{k}|} \mathbf{k} \times \{ \mathbf{k} \times [-\cos(kx + \varphi + \xi)\mathbf{a} + \sin(kx + \varphi + \xi)\mathbf{b}] \},$$
(15)
$$\mathbf{B} = \mathbf{k} \times [\cos(kx + \varphi)\mathbf{a} - \sin(kx + \varphi)\mathbf{b}]$$

where $\xi = \arccos(k^0/|\mathbf{k}|), |\mathbf{k}| > \alpha, k^0 = \sqrt{|\mathbf{k}|^2 - \alpha^2}$. Evidently, we can choose $\mathbf{a} \perp \mathbf{k}$ and $\mathbf{b} \perp \mathbf{k}$. Therefore in the PF

$$-\frac{1}{4}F\hat{F} = \mathbf{E} \cdot \mathbf{B} = \pm \alpha |\mathbf{a}| |\mathbf{b}| |\mathbf{k}|$$
(16)

and

$$\frac{1}{2}F^2 = \mathbf{B}^2 - \mathbf{E}^2 = \alpha(\mathbf{a}^2 - \mathbf{b}^2)|\mathbf{k}|\sin\left(2kx + 2\varphi + \xi\right)$$
(17)

Therefore, contrary to the massless case, $F\hat{F}$ and F^2 cannot vanish simultaneously except in the case $\mathbf{E} = \mathbf{B} = 0$. However, both \mathbf{E} and \mathbf{B} are perpendicular to \mathbf{k} so the wave front propagates along \mathbf{k} . Moreover, the angle between \mathbf{E} and \mathbf{B} is constant in time. The linear polarization is obtained for $\mathbf{a} = 0$ or $\mathbf{b} = 0$; in this case $\mathbf{E} \perp \mathbf{B}$. The elliptical polarization is obtained for \mathbf{a} and \mathbf{b} simultaneously different from zero; in this case $\mathbf{E} \cdot \mathbf{B} \neq 0$. Notice that for α going to zero we obtain the standard vacuum solution.

Now, the group velocity of the electromagnetic wave (15) is superluminal

$$\mathbf{v}_g = \nabla_k \omega(\mathbf{k}) = \left(\frac{\sqrt{k^{02} + \alpha^2}}{k^0}\right) \frac{\mathbf{k}}{|\mathbf{k}|}$$
(18)

while the phase propagates subluminally

$$\mathbf{v}_{\rm ph} = \left(\frac{k^0}{\sqrt{k^{02} + \alpha^2}}\right) \frac{\mathbf{k}}{|\mathbf{k}|}.$$
(19)

A very important question is the energy transport associated with the electromagnetic wave. The locally conserved canonical energy-momentum tensor, derived from the Lagrangian (10), is of the form

$$T^{\nu}_{\mu} = \frac{1}{4} \delta^{\nu}_{\mu} F^2 - F_{\nu\lambda} F^{\mu\lambda} - \alpha F_{\nu\lambda} A^{\lambda} u^{\mu}.$$
⁽²⁰⁾

It is evidently neither gauge-invariant, nor is T^{ν}_{μ} symmetrical in μ and ν . While this second deficiency is not serious, the first one is very unpleasant and the question of how to remedy this problem is unclear because the standard procedure fails in this case. However the field four-momentum

$$P_{\mu} := \int_{t=\text{const}} d\sigma^{\nu} T_{\nu\mu} = \int d^3 \mathbf{x} T_{0\mu}$$
(21)

is gauge-invariant. We can define the covariant four-momentum per volume as well as the gauge-invariant average density

$$p^{\mu} := \lim_{V \to 0} \frac{1}{V} \int_{V} d^{3} \mathbf{x} T_{0}^{\mu}$$
(22)

Now, for the monochromatic plane wave (15) in the PF, Eq. (22) leads to

$$p^{0} = \frac{(k^{0})^{2}}{2}(\mathbf{a}^{2} + \mathbf{b}^{2}),$$

$$\mathbf{p} = \frac{\mathbf{k}k^{0}}{2}(\mathbf{a}^{2} + \mathbf{b}^{2}).$$
(23)

Thus, in the PF

$$(p^{0})^{2} - \mathbf{p}^{2} = -\alpha^{2} (k^{0})^{2} \frac{(\mathbf{a}^{2} + \mathbf{b}^{2})^{2}}{4} \le 0$$
(24)

i.e., the energy transport is superluminal in this case also. Of course, the statements resulting from (18), (19) and (24) are true in all inertial frames, by Lorentz covariance.

3. Conclusions

Our discussion shows that a possible influence of the quantum preferred frame on an appropriate medium can cause tachyonic-like propagation for electromagnetic waves. It is interesting that solutions for the effective Maxwell equations (6) and (7) are very regular and similar to the usual ones. Therefore, this model appears to offer an alternative to standard proposals for explaining superluminal phenomena.

Acknowledgment

I acknowledge discussion with Piotr Kosiński and Wacław Tybor. This paper was financially supported by the University of Lodz.

References

- [1] A. M. Steinberg, P. G. Kwiat, and R. Z. Chiao, Phys. Rev. Lett. 71, 708 (1993).
- [2] A. M. Steinberg and R. Z. Chiao, Phys. Rev. A 51, 3525 (1995).
- [3] G. Nimtz, Gen. Rel. Grav. 31, 737 (1999).
- [4] G. Nimtz, Eur. Phys. J. B 7, 523 (1999).
- [5] L. J. Wang, A. Kuzmich, and A. Dogariu, Nature 406, 277 (2000).
- [6] Y. Aharonov and D. Z. Albert, Phys. Rev. D 24, 359 (1981).
- [7] A. Peres, Phys. Rev. A 61, 022117 (2000).
- [8] A. Peres, Quantum Theory: Concepts and Methods, (Kluwer, Dordrecht, 1995).
- [9] L. Hardy, Phys. Rev. Lett. 68, 2981 (1992).
- [10] I. C. Percival, Proc. R. Soc. London Ser. A 456, 25 (2000).
- [11] I. C. Percival, Phys. Lett. A 244, 495 (1998).
- [12] P. Caban and J. Rembieliński, Phys. Rev. A 59, 4187 (1999).
- [13] A. Suarez, quant-ph/0006053 (unpublished).
- [14] V. Scarani, W. Tittel, H. Zbinden, and N. Gisin, Phys. Lett. A 276, 1 (2000).
- [15] J. Rembieliński, Phys. Lett. A 78, 33 (1980).
- [16] J. Rembieliński, Int. J. Mod. Phys. 12, 1677 (1997).
- [17] S. Coleman and S. L. Glashow, hep-ph/9808446 (unpublished).
- [18] D. Colladay and V. A. Kostelecký, Phys. Rev. D 58, 11600 (1998).
- [19] R. Anderson, I. Vetharaniam, and G. E. Stedman, Phys. Rep. 295, 93 (1998).
- [20] D. E. Groom et al. (Particle Data Group), Eur. Phys. J. C 15, 1 (2000).
- [21] C. H. Lineweaver et al., Astrophys. J. 470, 28 (1996).

Sub– and Superluminal Velocities in Space with Vector Time

V. S. Barashenkov and M. Z. Yuriev Joint Institute for Nuclear Research, Laboratory of Computing Techniques and Automation, 141980 Dubna, Russian Federation

Within the bounds of the known relativistic theory the hypothesis of superluminal velocities allows one to influence the past, which leads to acausal paradoxes. We should like to stress, however, that this conclusion is based on the contradictory extension of the customary Lorentz transformations beyond the light barrier. Since at present no other prohibitions for faster-than-light signals carrying energy and information are known, the answer to the question does exist: such signals may or may not be obtained only from an experiment or from a more general theory. A generalization of a theory with vector time is considered, which allows some superluminal phenomena compatible with the principles of relativity and causality. Spreading of signats in the multitime world is characterized by peculiarities which can be used for an experimental determination of the time dimensionality of our world.

Keywords: superluminal

1. Is the hypothesis of superluminal speeds at variance with the experiment?

Let us consider the Lorentz transformation of a time interval Δt between two events separated by a space interval Δx :

$$\Delta t' = (\Delta t - \Delta x \cdot u/c^2)\gamma = \Delta t (1 - uv/c^2)\gamma < 0, \tag{1}$$

if the product of the moving body's speed $v = \Delta x / \Delta t$ and the relative velocity of the reference frame *u* exceeds unity, *i.e.*, $uv/c^2 > 1$ (*u* can be still smaller than *c* and the factor $\gamma = \sqrt{1 - u^2/c^2}$ well defined). The possibility of turning back the flow of time by considering the sequence of events from a moving body leads to difficulties of two types:

• Acausal phenomena contradicting our ideas about the time order of events appear when, for example, a bullet flies not from a hunter's gun to a targetcrow but, on the contrary, the crow emits the bullet and it runs back up the gun barrel. • Using superluminal signals one can change the past. In particular, an effect can destroy its cause: *e.g.*, by a faster-than-light signal we can prevent our birth or kill ourselves in the cradle and then the fact of our existence becomes an unexplicable puzzle.

At present there are two main viewpoints of this difficulty. Some authors (*e.g.*, E. Recami, see his review [1]) consider the phenomena with time reversals as really observable but apparently illusory events for which one can always find a genuine cause, just as we do when we hear the roar of a supersonic jet. However, this cannot explain or forbid suicide in the cradle, since it is not apparent, but can be actually be accomplished by a faster-than-light ray.

Another point of view shared by the majority of physicists (see the review [2] where a more detailed bibliography can be found) considers the difficulties as a proof of the obvious contradictoriness of the superluminal hypothesis and generally rejects the existence of superluminal signals carrying energy and information. Though we also share the latter opinion, it nevertheless appears to be insufficiently grounded. Indeed, as mentioned above, time reversal occurs, even if events are observed from a subluminal reference frame (e.q., from a conventional bicycle!). The existence of bodies with v > c assumes the possibility of using them as superluminal reference frames (*i.e.*, with u > c). A consistent generalization of Lorentz transformation in four-dimensional space-time, as proved in paper [15], is impossible ¹. The set of the equalities (1) is obviously true up to the last relation when uv > 1 is assumed. In four-dimensional space-time $\mathbf{x} = (x_1, x_2, x_3, ct)^T$. This, as it was shown in paper [15], at once turns the Lorentz group into an equivalent group of linear transformations $x'_{\mu} = \Lambda(v)_{\mu\nu}x^{\nu}$, with $Det\Lambda = \pm 1$. Successive use of several sub- and superluminal Lorentz transformations results in some symmetries which do not exist un our world — in a space dilation $\mathbf{x} \to \lambda \mathbf{x}$, in the time inversion $t \to -t$, etc. This means that the relations (1) at uv > 1 are not reliable and conclusions based on them are doubtful.

Of course, no superluminal phenomena carrying energy have been observed as yet. However, these results are related to the region of the phenomena described by known physics, and one cannot exclude the existence of inaccessible regions of events, outside the known ones, with in principle new laws where information can be carried with a faster-than-light speed without any violation of relativity and causality. One must also take into account that superluminal objects appear in various string models, in theories with high-order Lagrangians, by supersymmetrical generalizations *etc.*, and one may suspect that this fact is not only a disappointing theoretical

¹One should note that this difficulty is present in any theory with non-local interaction. For example, in a field theory with a form-factor where space- and time-like points enter into the interaction term $\int \phi(x_1)\phi(x_2)A(x_3)d^4x_1x_2x_3$ quite equivalently the reference frames tied to these points can be both types – sub- and superluminal. A formal relativistic invariant form of equations by itself does not provide complete Lorentz invariance of the theory.

failing, but is a reflection of some reality²

To answer the question as to the existence of faster-than-light motions, one must go into regions of unstudied phenomena where one can develop a consistent theory of relativity with velocities v > c.

2. Multitime velocity

In this respect interesting possibilities are provided by the theory of multidimensional time. Taking into account the apparent tendency of a symmetrization of physical theory with respect to the space and time co-ordinates, we assume that our world has the six-dimensional space-time structure

$$\hat{\mathbf{x}} = (\mathbf{x}, \hat{t})^T = (x_1, x_2, x_3, t_1, t_2, t_3)^T.$$
 (2)

(In what follows the tree-dimensional vectors in x- and t-subspaces will be denoted, respectively, by bold symbols and by a "hat", six-dimensional vectors will be marked, accordingly, by bold symbols with a hat).

The six-dimensional velocity vector is defined now as

$$\hat{\mathbf{v}} = \frac{\mathbf{d}\hat{\mathbf{x}}}{\mathbf{d}\hat{\tau}} = (\hat{\tau}\hat{\nabla})\hat{\mathbf{x}} = \tau_{\mathbf{i}}\frac{\partial\hat{\mathbf{x}}}{\partial\mathbf{t}_{\mathbf{i}}} = \frac{\mathbf{d}\hat{\mathbf{x}}}{\mathbf{d}\mathbf{t}} = (\mathbf{v}, \mathbf{c}\hat{\tau})^{\mathrm{T}},\tag{3}$$

where $\hat{\nabla} = (\partial/\partial t_1, \partial/\partial t_2, \partial/\partial t_3)$ and the unity vector $\hat{\tau} = d\hat{t}/dt$ with proper time t along the considered time trajectory.

If we notice that a differential of the squared length in the six-dimensional spacetime

$$ds^{2} = c^{2}(d\hat{t})^{2} - (d\mathbf{x})^{2} = c^{2}(dt)^{2} \left[1 - c^{-2}(d\mathbf{x}/dt)^{2}\right] = dt^{2}c^{2}/\gamma^{2}, \qquad (4)$$

where $\gamma = [1 - (v/c)^2]^{1/2}$, then the velocity vector can be written in the covariant form

$$\hat{\mathbf{u}} = d\hat{\mathbf{x}}/ds = (\gamma/c)d\hat{\mathbf{x}}/dt = \gamma\hat{\mathbf{v}}/c.$$
(5)

As in the customary onetime case the scalar product

$$\hat{\mathbf{u}}^2 = \gamma^2 \hat{\mathbf{v}}^2 / c^2 = \gamma^2 (c^2 \hat{\tau}^2 - \mathbf{v}^2 / c^2) = 1$$
(6)

and a light wave front always has a spherical form:

²In paper [16] the superluminal solutions for the Maxwell equations were discovered. Such solutions can be interpreted as describing "phase phenomena" which do not carry any information, like a bundle of sunbeams in a mirror. If we suppose that these solutions describe transportation of energy, then superluminal co-ordinate frames can be tied with bundles of such rays and the above mentioned difficulties appear. The discovered solutions can describe information carrying signals in a space-time with the dimensionalities $N > 3 \bigoplus 1$.



Fig. 1. An observer moves along the axis t_1 . From his viewpoint the speed of the body can exceed the light velocity.

$$\sum_{i} (\Delta x_{i}^{2} - c^{2} \Delta t_{i}^{2}) = \Delta t^{2} \sum_{i} (v_{i}^{2} - c^{2} \tau_{i}^{2}) = \Delta t^{2} (v^{2} - c^{2}) = 0,$$
(7)

i.e., in any direction of the x-subspace the body's speed does not exceed light velocity. Nevertheless, in a multitime world we can observe faster-than-light speeds of bodies.

3. Superluminal velocities

It is very important to emphasize that the body's speed v is defined with respect to an increment Δt along the body's time trajectory \hat{t} . If it is unknown and an observer uses instead of Δt his own proper time $\Delta t_p = \Delta t \cos \theta$ where θ is the angle between the body and observer's time trajectories, then the "speed" $v_p \equiv \Delta x/\Delta t_p = v/\cos\theta$ defined in this way may turn out to be larger than the light velocity. In this case the considered body behaves, from the observer's viewpoint, like a tachyon. For example, if $\theta \simeq \pi/2$, it passes any finite distance practically instantaneously and "grows old" straight away. However, as it was shown in the papers [5]-[7], Lorentz transformations depend on v but not on v_p ; therefore, in the multitemporal world no accausal effects can be observed by transformations to moving reference frames, contrary to true tachyons which transfer information to the new frame, judged by the observer, backwards in time (if the relations (1) are correct [2]). Superluminal velocities can also be observed in a more general case when the observer's time trajectory is, like a body, inclined with respect to the t_1 axis.

At the same time one should take into account that, as the onetime world with parallel trajectories $\hat{\tau}(t)$ is a particular case of the multitime world, the forbidding theorem on the superluminal generalization of the Lorentz transformations proved in paper [1] is also valid.

The discovery of any superluminal motions in experiments would be a serious indication of the multidimensionality of world time. As is known, faster-than-light objects are indeed observed by astronomers. Though up to now they have succeeded



Fig. 2. The creation of a component with the energy $\hat{E}' = \hat{\tau}' E \ge 0$ is accompanied, without fail, by the creation of a compensating component moving back in time with the energy $\hat{E}'' = \hat{\tau}'' E \le 0$. The energy vector is parallel to the time vector: $\hat{E} = E\hat{\tau}$.

in interpreting such phenomena within the limits of onetime notions as optical illusions (see, *e.g.*, [8, 9] where there are more detailed references), one cannot exclude that among such "superluminal objects" there are bodies moving along distinct time directions. We need more experimental information to identify such a possibility.

However, one must bear in mind that the creation of an object moving along time trajectories different from ours is possible only in exceptional cases when the known energy conservation law has vanished — in some cosmic cataclysm where new types of gravitation and electromagnetic waves can be produced or in very small space and time intervals (see Fig. 2). [10]- [12] ³.

Now let us consider some interesting peculiarities of signals spreading in the multitime world which can be used for an experimental determination of time dimensionality.

4. Detection of signals

As a simple example illustrating the peculiarities of the detection of signals in a multitime world, Cole and Starr considered a case when, under certain circumstances a splitting of time trajectories of a luminous body motionless in x-subspace and the observer occurs suddenly (Fig. 3) [13]. In the variant of a theory symmetrical with respect to every possible time direction considered by these authors, the light source gradually losing its energy (displacing into infrared region) remains visible some time after the moment of splitting. However, if time-reverse motions are forbidden (as is indeed observed in Nature), we come to quite a different conclusion. In particular, if the observer's time trajectory coincides with the t_1 axis, the luminous body becomes invisible at a given time because it occurs in the future with respect

³In paper[10] it was proposed that gravitational waves evolving along time trajectories different from ours could be detected by observing correlations of gravitation detector oscillations in two perpendicular directions. Another possibility for discovering motion along a distinct time trajectory may be based on the fact that the new components of the electromagnetic field created in the multitime world have longitudinal polarization and can be detected when the transversal components are excluded by any absorber.



Fig. 3. At a time t_o a splitting of time trajectories of an observer and a luminousbody $\hat{\tau}$ occurs. After that time $(t_1 > t_o)$ the body becomes invisible.

to the detector. The body can remain visible for some time after splitting only if the observer's trajectory has some inclination with respect to the t_1 axis.

A more complicated case is shown on the Fig. 4. One can see there that when the emitted light spreads in the plane (t_1, t_2) from the past to the future then the duration of observable luminescence is equal to

$$T \equiv (t_f - t_c) = (t_f - t_p) - R/c = \frac{R}{c} \left(\frac{\sin(\varphi + \theta)}{\sin \varphi} - 1 \right).$$
(8)

Here t_c is the time of the splitting, t_p is the observer's proper time when the light signal trajectory becomes parallel to the axis t_1 . At $t > t_f$ the time light signal propagates backward in time t_2 . R is a constant distance between the light source and the detector and φ is the angle between \hat{t} and $\hat{\tau}$. The inclination of the observer's trajectory with respect to the t_1 axis is denoted by θ .

If the time trajectory of a luminous body *intersects* the observer's trajectory (at the time $t = t_c$, see Fig. 4), the detector holds the light fixed in an interval from t_s when it fixes the ray emitted at a right angle to the t_1 axis up to the arrival time of the last visible signal t_f . For $t < t_s$ the body is too remote in the past and connection to it is possible only by means of subluminal signals (v < c). The rays emitted at $t > t_f$ cannot be observed by virtue of the causality principle. So, the duration of the visible light expressed throughout the observer's proper time is

$$T \equiv t_f - t_s = (t_f - t_c) + (t_c - t_s) = \frac{R}{c} \frac{\sin(\varphi + \theta)}{\sin\varphi} [1 + \cot(\varphi + \theta)].$$
(9)

As in the model considered by Cole and Starr [13, 14] the value of T is significant only for remote cosmic objects. For example, if R = 1 m and $\varphi = \theta = 1', 1^{\circ}, 40^{\circ}$, it is equal, respectively, to $2.10^{-5}, 4.10^{-7}, 10^{-8}c$. In a multitime world a large number of invisible time displaced bodies may be present around any observer. In this respect this world is much like a hypothetical tachyon theory world where there are also plenty of nonabsorbable objects [15]. One might imagine that an intersection of t-trajectories of the bodies between which a space distance is smaller than their dimensions must result in dramatic destruction of bodies. As such phenomena are



Fig. 4. At an observer's proper time t_c the luminous body's time trajectory $\hat{\tau}$ is split off from the observer's trajectory \hat{t} . The light is seen in the interval $t_s \div t_f$. Light spheres $(\hat{t} - \hat{\tau})^2 = (R/c)^2$ from which the observer can receive signals at different times *t* are dotted. The dotted lines with arrows show the trajectories of the first and last visible signals.

not observed in a our part of universe, this proves that time flow is single-directed in this region. The duration of the visible light from a light source *a moving* in *x*subspace depends on the value and direction of its velocity. However, qualitatively the picture remains the same as in the above considered static case. Particularly, if the observer's *t*-trajectory coincides with the t_1 axis and the light source moves in *x*-subspace with zero impact parameter (a head-on collision), then the luminescence becomes visible at a time

$$t_s = \frac{R_s}{c} \tan \varphi = \left(\frac{R_c}{c} + \beta t_s\right) \tan \varphi, \tag{10}$$

where $R_s = R(t_s)$ is the distance of the luminous body from the detector at the time t_s , R_c is the respective distance at the time when their t-trajectories intersect (t = 0), φ is the angle between these trajectories (Fig. 5A), $\beta = v/c$ is the relative velocity of the luminous body and the observer. Solving this equation, we obtain

$$t_s = \frac{R_c/c}{\tan \varphi - \beta}.$$
(11)

If at t = 0, the source and the detector come together and the velocity β is small ($\beta < \tan \varphi$, Fig. 2A), then the light is seen in an interval from t_s to $t_f = R_c/c$:

$$T = \frac{R_c}{c} [1 + 1/(\tan\varphi + \beta)].$$
(12)

By increasing the velocity $(\beta \ge \tan \varphi)$ we can stretch the time interval over all the left half-axis from $t_s = -\infty$ up to t_f . In the case when the light source moves away from the detector at time t = 0 and its velocity $\beta < \tan \varphi$ the light



Fig. 5. The bold tracks on the t_I axis are the intervals of visible light from a moving light source. The intersection of time trajectories is chosen as t = 0. The observer's light spheres are dotted. In case A the luminous body with velocity $\beta < \tan \varphi$ comes close to the observer at time t = 0. In case B the luminous body with hight velocity $\beta > \tan \varphi$ moves away from the observer at t = 0.

is observed, as before, in the interval from t_s to t_f . However, by $\beta \ge \tan \varphi$ (Fig. 5B) one more interval of the visible light beginning at $t_s = -\infty$ appears. The asymmetry of the cases of an approaching and receding light source is stipulated by detector asymmetry with respect to signals from the past and future.

5. Conclusion

In the limits of the commonly used superluminal generalizations of the Lorentz transformations, the hypothesis of faster-than-light velocities creates inadmissible paradoxes. However, this conclusion is doubtful since all the generalizations used are contradictory, and we cannot be fully confident that the basic relations (1) are correct. One cannot exclude the possibility that slices of reality exist where events developing with faster-than-light velocities and carrying energy can be observed. Is this statement right or wrong — it is now a question for experiment. Theories have been proposed, *e.g.*, multitime generalizations, which permit superlight processes without any violation of causality and relativity.

References

- [1] E. Recami, Riv. Nuovo Cim. 9 (1986) 1.
- [2] V. S. Barashenkov, M. Z. Yuriev, Hadr. J. 17 (1996) 470; JINR preprint E2-95-146, Dubna (1995).
- [3] L. Marchildon, A. E. Antipa, A. E. Everett, Phys. Rev. D 8 (1983) 1740.
- [4] W. A. Rodrigues, Jian-Yy Lu, Found. Phys. 37 (1997) 435.
- [5] E. A. B. Cole, S. A. Buchanan, J. Phys. A 15 (1982) L255.
- [6] E. A. B. Cole, Nuovo Cim. B 85 (1985) 105.
- [7] V. S. Barashenkov, Six-dimensional Spacetime Transformations, JINR preprint E2-97-83, Dubna (1997).

- [8] V. L. Ginzburg, Theoretical Physics and Astrophysics, Moccow Publ. House "Nauka" (1981).
- [9] I. I. Mirabel, L. F. Rodrigues, Nature 371 (1994) 46.
- [10] V. S. Barashenkov, A. B. Pestov, M. Z. Yuriev, Gen Rel. & Grav. 29 (1997) 1245.
- [11] V. S. Barashenkov, Found. Phys. 28 (1998) 471.
- [12] V. S. Barashenkov, M. Z. Yuriev, Quantum Field Theory with Three-dimensional Time Vector, JINR preprint P2-99-109, Dubna (1999).
- [13] E. A. B. Cole, J. M. Starr, Lett. Nuovo Cim. 43 (1985) 38.
- [14] E. A. B. Cole, J. M. Starr, Nuovo Cim. B. 105 (1990) 1091.
- [15] L. Basano, Lett. Nuovo Cim. 16 (1976) 562.

Superluminal Tunneling through Two Successive Barriers

Vladislav S. Olkhovsky Institute for Nuclear Research, Kiev-03028 Research Center "Vidhuk", Kiev, Ukraine Erasmo Recami Facoltà di Ingegneria, Università Statale di Bergamo, Dalmine (BG), Italy and I.N.F.N.–Sezione di Milano, Milan, Italy C.C.S., State University at Campinas, Campinas, S.P., Brazil Giovanni Salesi I.N.F.N.–Sezione di Milano, Milan, Italy Facoltà di Ingegneria, Università Statale di Bergamo, 24044– Dalmine (BG), Italy

We study the phenomenon of one-dimensional non-resonant tunneling through two successive (opaque) potential barriers, separated by an intermediate free region \mathcal{R} , by analyzing the relevant solutions to the Schroedinger equation. We find that the total traversal time does depends *not* only on the barrier widths (the so-called "Hartman effect"), but also on the \mathcal{R} width: so that the effective velocity in the region \mathcal{R} , between the two barriers, can be regarded as infinite. This agrees with the results known from the corresponding waveguide experiments, which simulated the tunneling experiment considered here due to the known formal identity between the Schroedinger and the Helmholtz equation. Finally, in an Appendix, we provide some general information (especially bibliographical) about the various sectors of science in which Superluminal motions seem to appear

Keywords: tunneling, tunneling time, superluminal

In this note we show that —when studying an experimental setup with two rectangular (opaque) potential barriers (Fig. 1)— the (total) tunneling phase time through the two barriers depends neither on the barrier widths *nor on the distance between the barriers*.

Let us consider the (quantum-mechanical) stationary solution for the one-dimensional (1D) tunneling of a non-relativistic particle, with mass m and kinetic energy $E = \hbar^2 k^2/2m = \frac{1}{2}mv^2$, through two equal rectangular barriers with height V_0 ($V_0 > E$) and width a, quantity $L - a \ge 0$ being the distance between them. The Schrödinger equation is

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + V(x)\psi(x) = E\psi(x), \qquad (1)$$

where V(x) is zero outside the barriers, while $V(x) = V_0$ inside the potential bar-



Fig. 1. The tunneling process, through two successive (opaque) potential barriers, considered in this paper. We show that the (total) tunneling phase time through the two barriers depends neither on the barrier widths *nor on the distance between the barriers*.

riers. In the various regions I $(x \le 0)$, II $(0 \le x \le a)$, III $(a \le x \le L)$, IV $(L \le x \le L + a)$ and V $(x \ge L + a)$, the stationary solutions to eq. (1) are the following

$$\psi_{\rm I} = e^{+ikx} + A_{\rm 1R} e^{-ikx},$$
 (2a)

$$\psi_{\rm II} = \alpha_1 \,\mathrm{e}^{-\chi x} + \beta_1 \,\mathrm{e}^{+\chi x},\tag{2b}$$

$$\psi_{\rm III} = A_{\rm 1T} \left[e^{i\kappa x} + A_{\rm 2R} e^{-i\kappa x} \right], \qquad (2c)$$

$$\psi_{\rm IV} = A_{\rm 1T} \left[\alpha_2 \, {\rm e}^{-\chi(x-L)} + \beta_2 \, {\rm e}^{+\chi(x-L)} \right], \tag{2d}$$

$$\psi_{\rm V} = A_{\rm 1T} A_{\rm 2T} \, {\rm e}^{ikx} \,, \tag{2e}$$

where $\chi \equiv \sqrt{2m(V_0 - E)}/\hbar$, and quantities A_{1R} , A_{2R} , A_{1T} , A_{2T} , α_1 , α_2 , β_1 and β_2 are the reflection amplitudes, the transmission amplitudes, and the coefficients of the "evanescent" (decreasing) and "anti-evanescent" (increasing) waves for barriers 1 and 2, respectively. Such quantities can be easily obtained from the matching (continuity) conditions:

$$\begin{cases} \psi_{\rm I}(0) = \psi_{\rm II}(0) \\ \partial a/a \\$$

$$\left\{ \left. \frac{\partial \psi_{1}}{\partial x} \right|_{x=0} = \left. \frac{\partial \psi_{11}}{\partial x} \right|_{x=0}, \tag{3b}$$

$$\begin{pmatrix}
\psi_{\rm II}(a) = \psi_{\rm III}(a) \\
\psi_{\rm III}(a) = \psi_{\rm III}(a)
\end{cases}$$
(4a)

$$\left\{ \left. \frac{\partial \psi_{\mathrm{II}}}{\partial x} \right|_{x=a} = \left. \frac{\partial \psi_{\mathrm{III}}}{\partial x} \right|_{x=a}, \tag{4b}$$

$$\begin{pmatrix} \psi_{\rm III}(L) = \psi_{\rm IV}(L) \\ \end{pmatrix}$$
 (5a)

$$\left. \left. \left. \frac{\partial \psi_{\text{III}}}{\partial x} \right|_{x=L} = \left. \frac{\partial \psi_{\text{IV}}}{\partial x} \right|_{x=L}, \tag{5b}$$

$$\begin{cases} \psi_{\rm IV}(L+a) = \psi_{\rm V}(L+a) \\ \frac{\partial \psi_{\rm IV}}{\partial \psi_{\rm IV}} \end{bmatrix} \tag{6a}$$

$$\frac{\partial T}{\partial x}\Big|_{x=L+a} = \frac{\partial T}{\partial x}\Big|_{x=L+a}.$$
 (6b)

Equations (3-6) are eight equations for our eight unknowns (A_{1R} , A_{2R} , A_{1T} , A_{2T} , α_1 , α_2 , β_1 and β_2). First, let us obtain the four unknowns A_{2R} , A_{2T} , α_2 , β_2 from eqs. (5) and (6) in the case of *opaque* barriers, *i.e.*, when $\chi a \to \infty$:

$$\alpha_2 \longrightarrow e^{ikL} \frac{2ik}{ik - \chi},$$
(7a)

$$\beta_2 \longrightarrow e^{ikL-2\chi_a} \frac{-2ik(ik+\chi)}{(ik-\chi)^2},$$
(7b)

$$A_{2\mathrm{R}} \longrightarrow \mathrm{e}^{2ikL} \frac{ik + \chi}{ik - \chi},\tag{7c}$$

$$A_{2\mathrm{T}} \longrightarrow \mathrm{e}^{-\chi a} \mathrm{e}^{-ika} \frac{-4ik\chi}{(ik-\chi)^2},\tag{7d}$$

We may then obtain the other four unknowns A_{1R} , A_{1T} , α_1 , β_1 from eqs. (3) and (4), again in the case $\chi a \to \infty$; one finds for instance that:

$$A_{1T} = -e^{-2\chi a} \frac{4i\chi k}{(\chi - ik)^2} A,$$
(8)

where

$$A \equiv \frac{2\chi k}{2\chi k \, \cos k(L-a) \, + \, (\chi^2 - k^2) \, \sin k(L-a)} \tag{9}$$

turns out to be *real*; and where, it must be stressed,

$$\delta \equiv \arg\left(\frac{ik+\chi}{ik-\chi}\right)$$

is a quantity that *does not depend* on a or on L. This is enough to conclude that the phase tunneling time (see, for instance, refs. [1-3])

$$\tau^{\rm ph}_{\rm tun} \equiv \hbar \, \frac{\partial \arg \left[A_{\rm 1T} A_{\rm 2T} \mathrm{e}^{ik(L+a)} \right]}{\partial E} = \hbar \, \frac{\partial}{\partial E} \arg \left[\frac{-4ik\chi}{(ik-\chi)^2} \right], \qquad (10)$$

while depending on the energy of the tunneling particle, *does not depend* on L + a (since it is actually independent both of a and L).

This result *not only* confirms the so-called "Hartman effect" [1,3] for the two opaque barriers —*i.e.*, the independence of the tunneling time from the opaque barrier widths—, but it *also* extends the effect by implying the total tunneling time to be independent even of L (see Fig. 1): something that may be regarded as further evidence of the fact that quantum systems seem to behave as non-local. It is important to stress, however, that the previous result holds only for non-resonant (nr) tunneling: *i.e.*, for energies far from the resonances that arise in region III due to interference between forward and backward travelling waves (a phenomemon quite analogous to the Fabry-Pérot phenomenon in the case of classical waves). Otherwise it is known that the general expression for (any) time delay τ near a resonance at $E_{\rm r}$ with half-width Γ would be $\tau = \hbar \Gamma [(E - E_{\rm r})^2 + \Gamma^2]^{-1} + \tau_{\rm nr}$.
The independence of tunneling-time from the width (a) of each one of the two opaque barriers is itself a generalization of the Hartman effect, and may be *a priori* understood —following refs. [4,5]— on the basis of the reshaping phenomenon which takes place inside a barrier.

The even more interesting tunneling-time independence from the distance L - a between the two barriers, can be understood on the basis of the interference between the waves leaving the first barrier (region II) and traveling in region III and the waves reflected from the second barrier (region IV) back into the same region III. Such interference has been shown [3] to cause an "advancement" (*i.e.*, an effective acceleration) of the incoming waves, a phenomenon similar to the analogous advancement expected even in region I. Using wavepacket language, we noticed in ref. [3] that the arriving wavepacket does interfere with the reflected waves that start to be generated as soon as the packet's forward tail reaches the (first) barrier edge: in this way (already before the barrier) the backward tail of the initial wavepacket decreases —for destructive interference with those reflected waves— to a larger degree than the forward one. This simulates an increase of the average speed of the entering packet: hence, the effective (average) flight-time of the approaching packet from the source to the barrier does decrease.

Consequently, the phenomena of reshaping and "advancement" (inside the barriers and to the left of the barriers) can qualitatively explain why the tunneling-time is independent of the barrier widths and of the distance between the two barriers. It remains impressive, nevertheless, that in region III —where no potential barrier is present, the current is non-zero and the wavefunction is oscillatory— the effective speed (or group-velocity) is practically *infinite*. Loosely speaking, one might say that the considered two-barrier setup is an "(intermediate) space destroyer". After some straightforward but rather bulky calculations, one can, moreover, see that the same effects (*i.e.*, the independence from the barrier widths and from the distances between the barriers) are still valid for any number of barriers, with different widths and different distances between them.

Finally, let us mention that the known similarity between photon and (nonrelativistic) particle tunneling [3-7] implies that our previous results hold also for photon tunneling through successive "barriers": for example, for photons in the presence of two successive band gap filters, such as two suitable gratings or two photonic crystals. Experiments should be easily realizable; while indirect experimental evidence seems to come from papers like [8].

At the classical limit, the (stationary) Helmholtz equation for an electromagnetic wavepacket in a waveguide is known to be mathematically identical to the (stationary) Schroedinger equation for a potential barrier;* so that, for instance, the tun-

^{*}These equations are, however, different (due to the different order of the time derivative) in the timedependent case. Nevertheless, it can be shown that they still have in common classes of analogous solutions, differing only in their spreading properties [3,7].

neling of a particle through and under a barrier can be simulated [3-7,9-11] by the traveling of evanescent waves along an undersized waveguide. Therefore, the results of this paper are to be valid also for electromagnetic wave propagation along waveguides with a succession of undersized segments (the "barriers") and of normalsized segments. This agrees with calculations performed, within the classical realm, directly from the Maxwell equations[10,11], and has already been confirmed by a series of "tunneling" experiments with microwaves: see refs.[9] and particularly [12].

Acknowledgements

Work partially supported by INFN, MURST and CNR (Italy), by CAPES (Brazil), and by I.N.R. (Kiev) and S.K.S.T. of Ukraine.

One of the authors (ER) is very grateful to Edward Kapuścik, Andrzej Horzela, Vladislav S.Olkhovsky and Jacek Jakiel for kind invitation to contribute to this Volume. The authors are grateful to S. Esposito, as well as to A. Agresti, D. Ahluwalia, G. G. N. Angilella, P. Bandyopadhyay, F. Bassani, R. Bonifacio, L. Bosi, C. Becchi, G.C. Cavalleri, R.Y. Chiao, R. Colombi, G. Degli Antoni, P. Falsaperla, F. Fontana, R. Garavaglia, A. Gigli, Berzolari, H. E. Hernández F., A. Horzela, J. Jakiel, E. Kapuścik, L. C. Kretly, G. Kurizki, J. León, G. Marchesini, R. Mignani, D. Mugnai, R. Petronzio, M. Pignanelli, G. Privitera, A. Ranfagni, A. Salanti, A. Shaarawi, A. Steinberg, J. W. Swart and M. Zamboni-Rached for scientific collaboration or stimulating discussions. For the active cooperation special thanks are due to G. Giuffrida, R. Salesi and M. T. Vasconselos.

References

- [1] T. E. Hartman: J. Appl. Phys. 33 (1962) 3427; J. R. Fletcher: J. Phys. C 18 (1985) L55.
- [2] E. H. Hauge and J. A. Støvneng: Rev. Mod. Phys. 61 (1989) 917.
- [3] V. S. Olkhovsky and E. Recami: *Phys. Reports* **214** (1992) 340; V. S. Olkhovsky, E. Recami, F. Raciti and A. K. Zaichenko: *J. Phys.I (France)* **5** (1995) 1351. See also V. S. Olkhovsky, E. Recami and J. Jakiel: "A new time-analysis of photon and (non-relativistic) particle tunneling" (to be subm. for pub.); D. Mugnai *et al.: Phys. Lett. A* **209** (1995) 227.
- [4] R. Y. Chiao, P. G. Kwiat and A. M. Steinberg: *Physica B* 175 (1991) 257. See also Th. Martin and R. Landauer: *Phys. Rev. A* 45 (1992) 2611; A. Ranfagni, D. Mugnai, P. Fabeni and G. P. Pazzi: *Appl. Phys. Lett.* 58 (1991) 774.
- [5] A. M. Steinberg, P. G. Kwiat and R. Y. Chiao: *Phys. Rev. Lett.* **71** (1993) 708; R. Y. Chiao, P. G. Kwiat and A. M. Steinberg: *Scient. Am.* **269** (1993), issue no.2, p.38.
- [6] Y. Japha and G. Kurizki: *Phys. Rev. A* 53 (1996) 586; "Causality and interference in electron and photon tunneling", preprint (Weizmann Inst. Sc.; Rehovot, Jan.1996); "Causality and non-locality in 'faster-than-light' photon travel", preprint (Weizmann Inst. Sc.; Rehovot, Sept.1996).
- [7] V. S. Olkhovsky and A. Agresti: in *Tunneling and its Applications* (World Scient.; Singapore, 1997), pp.327-355; J. Jakiel, V. S. Olkhovsky and E. Recami: *Phys. Lett.* A248 (1998) 156; A. M. Shaarawi and I. M. Besieris: *J. Phys. A: Math. Gen.* 33 (2000) 1.
- [8] S. Chu and S. Wong: *Phys. Rev. Lett.* 48 (1982) 738; see also C. G. B. Garret and D. E. McCumber: *Phys. Rev. A* 1 (1970) 305; L. J. Wang, A. Kuzmich and A. Dogariu: Nature 406 (2000) 277.
- [9] A. Enders and G. Nimtz: J. Phys.I (France) 3 (1993) 1089; A.Ranfagni, P. Fabeni, G. P. Pazzi and D. Mugnai: Phys. Rev. E 48 (1993) 1453; Ch. Spielmann, R. Szipocs, A. Stingl and F. Krausz: Phys. Rev. Lett. 73 (1994) 2308, Ph. Balcou and L. Dutriaux: Phys. Rev. Lett. 78 (1997) 851; V.Laude and P. Tournois: J. Opt. Soc. Am. B 16 (1999) 194.

- [10] A. P. L. Barbero, H. E. Hernández F., and E. Recami: "On the propagation speed of evanescent modes" [Lanl Archives # physics/ 9811001], to appear in Phys. Rev. E. Cf. also E. Recami, H. E. Hernández F., and A. P. L. Barbero: Ann. der Phys.7 (1998) 764; H. M. Brodowsky, W. Heitmann and G. Nimtz: Phys. Lett. A 222 (1996) 125.
- [11] M. Zamboni-Rached, E. Recami and F. Fontana: "Localized Superluminal Solutions to Maxwell Equations propagating along a normal-sized waveguide", Lanl Archives # physics/0001039, subm. for pub.; cf. also E. Recami, F. Fontana and R. Garavaglia: *Int. J. Mod. Phys. A* 15 (2000) 2793; E. Recami: in *Causality and Locality in Modern Physics*, ed. by G. Hunter, S. Jeffers and J.-P. Vigier (Kluwer; Dordrecht, 1998), p.113.
- [12] G. Nimtz, A. Enders and H. Spieker: in *Wave and Particle in Light and Matter*, ed. by A. van der Merwe & A. Garuccio (Plenum; New York, 1993); *J. de Physique I* 4 (1994) 565; A. Enders and G. Nimtz: *Phys. Rev. B* 47 (1993) 9605.

A P P E N D I X Some information about the experimental sectors of science in which Superluminal motions seem to appear

Introduction

The question of Super-luminal $(V^2 > c^2)$ objects or waves [tachyons: a term coined by G. Feinberg] has a long story, starting perhaps with Lucretius' De Rerum Natura (cf., book 4, line 201). Still in pre-relativistic times, we may recall e. g., the papers by A. Sommerfeld (quoted in refs. [A1, A2]). In relativistic times, our problem again attracted attention essentially in the fifties and sixties, in particular after the papers by E. C. George Sudarshan etal., and later on by E. Recami, R. Mignani, et al. [who coined the term bradyons for slower-than-light objects, and brought the expressions subluminal and superluminal into popular use through their works at the beginning of the seventies], as well as by H. C. Corben and others (to confine ourselves to the theoretical research). For references, one can check pages 162-178 in ref. [A1], where about 600 citations are listed; pages 285-290 in ref. [A3]; pages 592-597 of ref. [A4] or pages 295-298 of ref. [A5]; as well as the large bibliographies by V. F. Perepelitsa[A6] and as the book in ref. [A7]. In particular, for the causality problems one can see refs. [A1,A8] and references therein, while for a model theory for tachyons in two dimensions one can consult refs. [A1,A9]. The first experiments to seek tachyons were performed by T. Alväger et al. .; citations about the early experimental quest for superluminal objects may be found, e. g., in refs. [A1,A10].

The subject of tachyons is now back in fashion, especially because of the fact that at least *four* different experimental sectors of physics seem to indicate the existence of Superluminal objects [the old habit introducted by Mignani and Recami of writing Superluminal with a capital S]. We wish to provide in the following some information (mainly bibliographical) about the experimental results obtained in each one of these 4 different sectors of physics.

FIRST: Negative Mass-Square Neutrinos

Since 1971 it has been known that the experimental square-mass of muon - neutrinos was negative (with low statistical significance, but systematically). If confirmed, this would correspond (within the ordinary, naïve approach to relativistic particles) to an imaginary mass and, therefore, to a Superluminal speed; in a non-naïve approach[A1], *i.e.*, within a Special Relativity theory extended to include tachyons [Extended Relativity (ER)], the free tachyon "dispersion relation" becomes $E^2 - p^2 = -m_0^2$. See e. g. E. V. Shrum & K. O. H. Ziock: *Phys. Lett. B* **37** (1971) 114; D. C. Lu *et al.* : *Phys. Rev. Lett.* **45** (1980) 1066; G. Backenstoss *et al.* : *Phys. Lett. B* **43** (1973) 539; H. B. Anderhub *et al.* : *Phys. Lett. B* **114** (1982) 76; R. Abela *et al.* : *Phys. Lett. B* **146** (1984) 431; B. Jeckelmann *et al.* : *Phys. Rev. Lett.* **56** (1986) 1444.

For the theoretical point of view, see

E. Giannetto, G. D. Maccarrone, R. Mignani & E. Recami: Phys. Lett. B 178 (1986)

115-120, and references therein; and S. Giani's work; see also E. Recami: "Classical tachyons and possible applications", *Rivista Nuovo Cim.* **9** (1986), issue no. 6, and references therein.

Recent experiments showed that electron-neutrinos also have negative masssquare. See e. g. R. G. H. Robertson *et al.* : *Phys. Rev. Lett.* **67** (1991) 957; A. Burrows *et al.* : *Phys. Rev. Lett.* **68** (1992) 3834; Ch. Weinheimer *et al.* : *Phys. Lett. B* **300** (1993) 210; E. Holtzshuh *et al.* : *Phys. Lett. B* **287** (1992) 381; H. Kawakami *et al.* : *Phys. Lett. B* **256** (1991) 105, and so on. See also the reviews or comments by M. Baldo Ceolin: "Review of neutrino physics," invited talk at the *XXIII Int. Symp. on Multiparticle Dynamics* (Aspen, CO; Sept. 1993)"; E. W. Otten: *Nucl. Phys. News* **5** (1995) 11.

SECOND: Galactic "Micro-Quasars", etc. (Apparent Superluminal expansions observed inside quasars, some galaxies, and –as discovered very recently– in some galactic objects, called "micro-quasars")

Since 1971 apparent Superluminal expansions have ben observed in many quasars (and even a few galaxies) [*Nature*, for instance, dedicated a cover to these observations on 2 Apr. 1981]. Such apparent Superluminal expansion was the consequence of the experimentally measured angular separation rates, once the (large) distance of the sources from the Earth was taken into account. From the experimental point of view, a quote from the book *Superluminal Radio Sources*, ed. by J. A. Zensus & T. J. Pearson (Cambridge Univ. Press; Cambridge, UK, 1987), and references therein, is sufficient.

The distance those "Superluminal sources", however, it is not well known; or, at least, the (large) distances usually adopted have been strongly criticized by H. Arp *et al.*, who maintain that quasars are much nearer objects: so that all the abovementioned data can no longer be easily used to infer (apparent) Superluminal motions. However, very recently, GALACTIC objects have been discovered, in which apparent Superluminal expansions occur; and the distances of galactic objects can be more precisely determined. From the experimental point of view, see the papers by I. F. Mirabel & L. F. Rodriguez. : "A superluminal source in the Galaxy," *Nature* **371** (1994) 46 [with a *Nature*'s comment, "A galactic speed record," by G. Gisler, on page 18 of the same issue]; and by S. J. Tingay *et al.* (20 authors): "Relativistic motion in a nearby bright X-ray source," *Nature* **374** (1995) 141.

From the theoretical point of view, both for quasars and "micro-quasars", see E. Recami, A. Castellino, G. D. Maccarrone & M. Rodonò: "Considerations about the apparent Superluminal expansions observed in astrophysics," *Nuovo Cimento* B 93 (1986) 119. See also E. Recami: ref. [A1], and cf. R. Mignani & E. Recami: Gen. Relat. Grav. 5 (1974) 615. In particular, let us recall that a single Superluminal source of light would be observed: (i) initially, in the phase of "optic boom" (analogous to the acoustic "boom" by an aircraft that travels with constant super-sonic speed) as an intense, suddenly-appearing source; (ii) later on,

as a source which splits into two objects receding one from the other with velocity v > 2c [see the quoted refs.].

THIRD: Tunneling photons = Evanescent waves

This is the sector that has attracted the most attention from the scientific and nonscientific *press*: it started in *Scientific American* in Aug. 1993 followed by *Nature* (comment "Light faster than light?" by R. Landauer) on Oct. 21, 1993; then, *New Scientist* (editorial "Faster than Einstein" on p. 3, plus an article by J. Brown on p. 26) in April 1995; and then *Newsweek* (19 June 1995, article by S. Begley, p. 44) and all the newspapers and magazines of the world (in Brazil, *e.g.*, the *Folha de São Paulo*, *etc.*; in Italy, e. g., *La Stampa*, *La Repubblica*, *Focus*, *Panorama*, *etc.*).

Evanescent waves were predicted by ER [cf., page 158 in ref. [A1], and refer ences therein] to be faster-than-light. Even more, they can be regarded as consisting of tunneling photons, due to the known methematical identity of the Schroedinger equation (in the presence of a potential barrier) and the Helmholtz equation (for waves travelling, e.g., down a waveguide): and it has been known for some time [cf.V. S. Olkhovsky & E. Recami: *Phys. Reports* **214** (1992) 339, and refs. therein] that tunneling wave packets can move with Superluminal group velocities inside the barrier. Therefore, due to the theoretical analogies between tunneling particles (e.g., electrons) and tunneling photons, it was expected also on the basis of Quantum Mechanics that evanescent waves could be Superluminal. This has actually been confirmed in a series of famous experiments.

The first experiments were performed at Cologne, Germany, by Guenter Nimtz et al., and published in 1992. Let us quote: A. Enders & G. Nimtz: J. de Physique-I 2 (1992) 1693; 3 (1993) 1089; Phys. Rev. B 47 (1993) 9605; Phys. Rev. E 48 (1993) 632; G. Nimtz, A. Enders & H. Spieker: J de Physique-I 4 (1994) 1379; W. Heitmann & G. Nimtz: Phys. Lett. A 196 (1994) 154; G. Nimtz: Physik Bl. 49 (1993) 1119; "New knowledge of tunneling from photonic experiments," in Tunneling and its Implications (World Scient. ; Singapore, in press); G. Nimtz & W. Heitmann: "Photonic bands and tunneling," in Advances in Quantum Phenomena, ed. by E. G. Beltrametti and J. -M. Lévy-Leblond (Plenum Press; New York, 1995), p. 185; Prog. Quant. Electr. 21 (1997) 81; G. Nimtz, A. Enders and H. Spieker: "Photonic tunneling experiments: Superluminal tunneling," in Wave and Particle in Light and Matter, ed. by A. van der Merwe & A. Garuccio (Plenum; New York, 1993); J. de Physique-I 4 (1994) 565; H. Aichmann & G. Nimtz: "Tunneling of a FM-Signal: Mozart 40," submitted for pub. These are important experimental papers. Nimtz et al. also made similar simulations by computer (on the basis of the Maxwell eqs.), reproducing the related experimental results, where they exist, accurately: cf. H. M. Brodowsky, W. Heitmann & G. Nimtz, Phys. Lett. A 222 (1996) 125-129.

Other famous experiments have been performed at Berkeley; their results appeared in 1993 in A. M. Steinberg, P. G. Kwiat & R. Y. Chiao: *Phys. Rev. Lett.* **71**

(1993) 708, and, simultaneously, in R. Y. Chiao, P. G. Kwiat & A. M. Steinberg: *Scientific American 269* (1993), issue no. 2, p. 38. *Cf.* also A. M. Steinberg & R. Y. Chiao: *Phys. Rev. A* **51** (1995) 3525; P. G. Kwiat *etal.*: *Phys. Rev. A* **48** (1993) R867; E. L. Bolda *et al.*: *Phys. Rev. A* **48** (1993) 3890.

Further experiments on Superluminal evanescent waves have been performed at Florence: see, *e.g.*, A. Ranfagni, P. Fabeni, G. P. Pazzi & D. Mugnai: *Phys. Rev. E* **48** (1993) 1453. The last experiments (as far as we know) were made at Vienna: Ch. Spielmann, R. Szipocs, A. Stingl & F. Krausz: *Phys. Rev. Lett.* **73** (1994) 2308, and at Rennes and Orsay: Ph. Balcou & L. Dutriaux: *Phys. Rev. Lett.* **78** (1997) 851; V. Laude & P. Tournois: *J. Opt. Soc. Am. B* **16** (1999) 194.

For the *theoretical* point of view, see the above-quoted V. S. Olkhovsky & E. Recami: *Phys. Reports* **214** (1992) 339, and refs. therein; and V. S. Olkhovsky, E. Recami, F. Raciti & A. K. Zaichenko: *J. de Physique-1* **5** (1995) 1351-1365. See also pages 158 and 116-117 of the already quoted ref. [A1]; D. Mugnai *et al.* : *Phys. Lett. A* **209** (1995) 227-234; E. Recami, F. Fontana & R. Garavaglia: *Int. J. Mod. Phys. A* **15** (2000) 2793-2812; and V. S. Olkhovsky, E. Recami & G. Salesi: Lanl Archives # quant-ph/0002022.

The most interesting experiment of this type seems to be the one performed with two "barriers" (for instance, with two segments of undersized waveguide separated by a normal waveguide); for suitable frequency-band pulses —*i.e.*, for non-resonant "tunneling" —, it has been found that total crossing time does *not* depend on the length of the intermediate (normal) waveguide: that is to say, the pulse speed along the latter is infinite[A11]. This agrees once more with the predictions of Quantum Mechanics for tunneling through two successive opaque barriers (the tunneling phase time does *not* depend on the distance between the barriers[A12]). Such an important experiment could and should be repeated, also taking advantage of the fact that evanescence regions can be easily constructed in many different ways or by different "photonic band-gap materials" and gratings (since one can use multilayer dielectric mirrors, semiconductors, photonic crystals, *etc.*)

At this point, let us observe *also* the following. Even if in ER all the ordinary causal paradoxes seem to be solvable[A1,A8], nevertheless, one ought to bear in mind that (whenever an object, \mathcal{O} , is encountered travelling at Superluminal speed) negative contributions should be expected to the tunneling times[A13]: and this ought not to be regarded as unphysical[A1,A8]. In fact, whenever an "object" \mathcal{O} *overcomes* the infinite speed with respect to a certain observer, it will afterwards appear to the same observer as its "*anti* object" \mathcal{O} travelling in the opposite *space* direction[A1,A8]. For instance, when passing from the lab to a frame \mathcal{F} moving in the *same* direction as the particles or waves entering the barrier region, the objects \mathcal{O} penetrating through the final part of the barrier (with almost infinite speed[A14]) will appear in the frame \mathcal{F} as anti-objects \mathcal{O} crossing that portion of the barrier *in the opposite space-direction*[A1,A8]. In the new frame \mathcal{F} , therefore, such antiobjects \overline{O} would yield a *negative* contribution to the tunneling time, which could even turn out, in total, to be negative. What we want to stress here is that the appearance of such negative times is predicted by Relativity itself, on the basis of the ordinary postulates[A1,A8,A13,A14]. From the theoretical point of view, besides refs. [A13,A14,A8,A1], see also R. Y. Chiao, A. E. Kozhekin A. E., and G. Kurizki: *Phys. Rev. Lett.* **77** (1996) 1254; C. G. B. Garret & D. E. McCumber: *Phys. Rev.* A **1** (1970) 305. From the (quite interesting!) experimental point of view, see S. Chu & Wong W. : *Phys. Rev. Lett.* **48** (1982) 738; M. W. Mitchell & R. Y. Chiao: *Phys. Lett. A* **230** (1997) 133-138; G. Nimtz: *Europ. Phys. J. B* (to appear as a Rapid Note); L. Wang *et al.* : *Nature* **406** (2000) 277; further experiments are being performed at Glasgow [D. Jaroszynski, private communication].

Finally, let us emphasize that faster-than-c propagation of light pulses can be (and was, in same cases) observed also by taking advantage of anomalous dispersion near an absorbing line, or nonlinear and linear gain lines, or nondispersive dielectric media, or inverted two-level media, as well as of some parametric processes in nonlinear optics (cf. G. Kurizki *et al.*).

FOURTH: Superluminal motions in Electrical and Acoustical Engineering – The "X-shaped waves"

This fourth sector is perhaps the most important one.

Starting with the pioneering work by H. Bateman, it gradually became known that all the (homogeneous) wave equations —in a general sense: scalar, electromagnetic and spinor — admit solutions with subluminal (v < c) group velocities [A15]. More recently, Super-luminal (V > c) solutions have also been constructed for those homogeneous wave equations, in refs. [A16] and quite independently in refs. [A17]: in some cases just by applying a Superluminal Lorentz "transformation" [A1,A18]. It has been also shown that the same happens even in the case of acoustic waves, with the presence in this case of "sub-sonic" and "Super-sonic" solutions [A19]. Particular attention has been attracted to the fact that some of the new solutions are "undistorted progressive waves" (namely, represent localized, non-diffractive waves). One can expect all such solutions to exist, *e.g.*, also for seismic wave equations. More intriguingly, one might expect the same to be true in the case of gravitational waves too.

It is interesting to remark that the Super-sonic and Super-luminal solutions put forward in refs. [A20] —some of them already experimentally realized [A21] appear to be (generally speaking) X-shaped, just as predicted in 1980–1982 by A. O. Barut, G. D. Maccarrone & E. Recami in ref. [A21]; so that they now have been preliminarily called "X-waves."

In this regard, from the theoretical point of view, we may cite pages 116-117, and pages 59 (fig. 19) and 141 (fig. 42), in E. Recami: ref. [A1]. Even more, see the abovementioned A. O. Barut, G. D. Maccarrone & E. Recami: "On the shape of tachyons," *Nuovo Cimento A* **71** (1982) 509-533 (and refs. [A21]) where "X-shaped

waves" are predicted and discussed; *cf.* also E. Recami [A20], which appeared in *Physica A*. From the quoted papers it is also clear why the X-shaped waves keep their form while travelling (non-dispersive waves): a property that has elicited high interest from electrical and acoustical engineering. New experimental and theoretical work is going on (*e.g.*, by F. Fontana *et al.* at the "Pirelli Cavi", Milan, Italy, with pulsed lasers; and by H. E. Hernandez F. *et al.* at the F. E. E. C. of Unicamp, Campinas, S. P.). Let us mention in particular work by P. Saari, H. Sõnajalg *et al.* at Tartu, Estonia (see, *e.g.*, *Opt. Lett.* **22** (1997) 310; *Laser Phys.* **7** (1997) 32), who *experimentally produced Superluminal X-shaped light waves*[A22] in optics, and work by D. Mugnai, A. Ranfagni and R. Ruggeri, who produced at IROE/CNR, Florence, Superluminal X-shaped beams in the realm of microwaves. [A22] Simultaneously, as expected on the basis of ER, also (non-truncated) X-shaped beams with finite total energy have been constructed[A23]; while many new Localized Superluminal Solutions to the Maxwell equations have been found (some of them generalizing the X-shaped beams)[A23].

Further (Numbered) References of the Appendix:

[A1] E. Recami: Rivista Nuovo Cim. 9 (1986), issue no. 6, 1.

[A2] P. O. Fröman: Arch. Hist. Exact Sci. 48 (1994) 373.

[A3] R. Mignani & E. Recami: Rivista Nuovo Cim. 4 (1974) 209; E398.

[A4] E. Recami: in *Albert Einstein 1879-1979: Relativity, Quanta and Cosmology*, ed. by F. De Finis & M. Pantaleo, vol. 2 (Johnson Reprint Co. ; New York, 1979), p. 537. This book exists also in Italian and in Russian.

[A5] P. Caldirola & E. Recami: in *Italian Studies in the Philosophy of Science*, ed. by M. L. Dalla Chiara (Reidel: Boston, 1980), p. 249.

[A6] V. F. Perepelitsa: Reports ITEF-100 and ITEF-165 (Institute of Theoretical and Experimental Physics; Moscow, 1980).

[A7] E. Recami (ed.): *Tachyons, Monopoles, and Related Topics* (North-Holland; Amsterdam, 1978).

[A8] E. Recami: Found. of Phys. 17 (1987) 239; Lett. N. Cim. 44 (1985) 587.

[A9] E. Recami & W. A. Rodrigues: in *Gravitational Radiation and Relativity*, ed. by J. Weber & T. M. Karade (World Scient.; Singapore, 1985), pp. 151.

[A10] See e. g. D. F. Bartlett *et al.* : *Phys. Rev. D* 18 (1978) 2253; P. N. Bhat *et al.* : *J. Phys. G* 5 (1979) L13. Cf. also A. S. Goldhaber & F. Smith: *Rep. Progr. Phys.* 38 (1975) 757; L. W. Jones: *Rev. Mod. Phys.* 49 (1977) 717.

[A11] G. Nimtz, A. Enders & H. Spieker: in *Wave and Particle in Light and Matter*, ed. by A. van der Merwe & A. Garuccio (Plenum; New York, 1993); [*J. de Physique-I* 4 (1994) 565. See also A. Enders & G. Nimtz: *Phys. Rev. B* 47 (1993) 9605.

[A12] V. S. Olkhovsky, E. Recami & G. Salesi: "Tunneling through two successive barriers and the Hartman (Superluminal) effect" [Lanl Archives # quant-ph/0002022], Report INFN/FM–00/20 (Frascati, 2000), submitted for pub.; S. Esposito: (in preparation). Cf. also A. M. Shaarawi & I. M. Besieris: *J. Phys. A* **33** (2000) 1.

[A13] See V. S. Olkhovsky, E. Recami, F. Raciti & A. K. Zaichenko: *J. de Physique I* **5** (1995) 1351. See also refs. [A1,A8] and E. Recami, F. Fontana & R. Garavaglia: *Int. J. Mod. Phys. A* **15** (2000) 2793.

[A14] V. S. Olkhovsky & E. Recami: *Phys. Reports* **214** (1992) 339, and refs. therein; V. S. Olkhovsky, E. Recami, F. Raciti & A. K. Zaichenko: *J. de Physique-I* **5** (1995) 1351; A. P. L. Barbero, H. E. Hernández F., & E. Recami: "On the propagation speed of evanescent modes" [LANL Archives # physics/

9811001], in press in *Phys. Rev. E*; V. S. Olkhovsky, E. Recarni & G. Salesi: ref. [A12]. See also E. Recarni, H. E. Hernández F. , & A. P. L. Barbero: *Ann. der Phys.* **7** (1998) 764; S. Esposito: (in preparation).

[A15] H. Bateman: *Electrical and Optical Wave Motion* (Cambridge Univ. Press; Cambridge, 1915). See also J. N. Brittingham: J. Appl. Phys. 54 (1983) 1179; R. W. Ziolkowski: J. Math. Phys. 26 (1985) 861; J. Durnin: J. Opt. Soc. 4 (1987) 651; A. O. Barut et al. : Phys. Lett. A 143 (1990) 349; Found. Phys. Lett. 3 (1990) 303; Found. Phys. 22 (1992) 1267.

[A16] A. O. Barut et al. : Phys. Lett. A 180 (1993) 5; 189 (1994) 277.

[A17] R. Donnelly & R. W. Ziolkowski: *Proc. Roy. Soc. London A* 440 (1993) 541; I. M. Besieris, A. M. Shaarawi & R. W. Ziolkowski: *J. Math. Phys.* 30 (1989) 1254; S. Esposito: *Phys. Lett. A* 225 (1997) 203; J. Vaz & W. A. Rodrigues; *Adv. Appl. Cliff. Alg.* S-7 (1997) 457. Cf. also R. W. Ziolkowski, D. K. Lewis & B. D. Cook: *Phys. Rev. Lett.* 62 (1989) 147. [A18] See ref. [A1] and refs. therein. See also E. Recami & W. A. Rodrigues: "A model theory for tachyons in two dimensions," in *Gravitational Radiation and Relativity*, ed. by J. Weber & T. M. Karade (World Scient.; Singapore, 1985), pp. 151.

[A19] J. -y. Lu and J. F. Greenleaf: *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 37 (1990) 438; 39 (1992) 19.

[A20] J. -Y. Lu & J. F. Greenleaf: *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **39** (1992) 441; R. W. Ziolkowski, I. A. Besieris & A. M. Shaarawi, *J. Opt. Soc. Am. A* **10** (1993) 75; E. Recami: *Physica A* **252** (1998) 586; J. -y. Lu, J. F. Greenleaf & E. Recami: "Limited diffraction solutions to Maxwell (and Schroedinger) equations" [Lant Arch. # physics/9610012], Report INFN/FM–96/01 (I. N. F. N. ; Frascati, Oct. 1996); M. Zamboni-Rached, E. Recami & F. Fontana: "Localized Superluminal solutions to Maxwell equations propagating along a normal-sized waveguide" [Lanl Archives # physics/0001039], subm. for pub.

[A21] A. O. Barut, G. D. Maccarrone & E. Recami: Nuovo Cimento A 71 (1982) 509. See also E. Recami & G. D. Maccarrone: Lett. Nuovo Cim. 28 (1980) 151; P. Caldirola, G. D. Maccarrone & E. Recami: Lett. Nuovo Cim. 29 (1980) 241.

[A22] P. Saari & K. Reivelt: "Evidence of X-shaped propagation-invariant localized light waves," *Phys. Rev. Lett.* **79** (1997) 4135; H. Sõnajalg, M. Rätsep & P. Saari: *Opt. Lett.* **22** (1997) 310. D. Mugnai, A. Ranfagni and R. Ruggeri: *Phys. Rev. Lett.* **84** (2000) 4830.

[A23] M. Zamboni-Rached, E. Recami & H. E. Hernández-Figueroa: (in preparation); I. Besieris, M. Abdel-Rahman, A. Shaarawi & A. Chatzipetros: "Two fundamental representations of localized pulse solutions to the scalar wave equation", *Progress in Electromagnetic Research (PIER)* **19** (1998) 1.

Tunneling in Two- and Three-layer Systems with Allowance for Dynamic Image Forces

A. M. Gabovich, V. M. Rozenbaum, and A. I. Voitenko Institute of Physics, National Academy of Sciences, Prospekt Nauki 46, MSP 03650 Kiev-22, Ukraine

A short review is given of the original treatment of the dynamic image forces and charge tunneling in two- and three-layer systems. Both metallic and semiconducting electrodes are studied. The linkage between plasma-like medium non-adiabatic response and the notion of the tunneling time is demonstrated.

Keywords: tunneling, dynamic image forces, dielectric permittivity, plasmon damping

1. Introduction

The problem of the tunneling time τ_{tun} in quantum mechanics has turned out to be extremely important and difficult [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11]. It emerged [12, 13] soon after the tunneling concept itself was introduced for electronic [14, 15, 16], atomic [17, 18], and nuclear physics [19, 20, 21, 22, 23], as well as for low-temperature chemistry [18, 24, 25, 26, 27]. Although far from being solved, the problem has, nevertheless, provided some insight into condensed matter physics [15]. Specifically, charged particles moving near interfaces and in thin interlayer gaps excite virtual or real collective oscillations of the metal (semiconductor) plasma [28, 29, 30, 31, 32, 33, 34]. In this case the polarization (image) forces differ from the classical ones due to dynamic (nonadiabatic) renormalization [35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45]. For sub-barrier (tunnel) processes, two time scales are inherent in the problem: τ_{tun} mentioned above and ω_s^{-1} , the inverse circular frequency of surface plasmons. The dynamic corrections are essential if $\tau_{tun}\omega_s < 1$, *i.e.*, when the electrode plasma response is retarded with respect to the projectile Coulomb field action [41, 43, 46, 47]. Otherwise, only static corrections of a different origin modify the classical result [32, 35, 39, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63].

In this article we summarize some of our recent results concerning dynamic image forces and the related topic of electron interelectrode tunneling. Our perturbationbased approach is set forth and justified by model calculations. It is shown that for small nonadiabatic corrections the explicit choice between different characteristic tunneling times can be avoided. However, implicitly this choice should be made



Fig.1. The charge moving according to the z(t) law in the three-layer system with dielectric permittivities $\varepsilon_i(\mathbf{k}, \omega)$.

while developing a future nonperturbative dynamic tunneling theory. The existing ones, although very complicated, are neither unambiguous nor self-consistent [40, 41, 42, 44, 45, 58, 59, 60, 64, 65, 66, 67, 68, 69, 70, 71, 72], so that the challenge to theoreticians still persists.

2. Formulation of the problem

The energy of classical image forces for the charge q near a flat vacuum–metal interface (the subscript "surf") has the form (see, *e.g.*, [39])

$$W_{\rm surf}^{\rm cl}(r) = -q^2/4r,\tag{1}$$

where r > 0 is the distance from the interface. At the same time, in the vacuum slab between two classical metallic electrodes the contributions of the infinite sequence of images converge into the following expression [73]:

$$W_{\rm slab}^{\rm cl}(z) = \frac{q^2}{8l} \left[2\ln\gamma + \psi\left(\frac{1}{2} - \frac{z}{2l}\right) + \psi\left(\frac{1}{2} + \frac{z}{2l}\right) \right].$$
 (2)

Here 2l is the slab width, $\psi(z)$ is the digamma function, $\gamma = 1.7810...$ is the Euler constant, the distance z is reckoned from the center of the interlayer (see Fig. 1). $W_{\text{slab}}^{\text{cl}}(z)$ diverges at the interfaces $z = \pm l$ in the same manner as $W_{\text{surf}}^{\text{cl}}(r \to 0)$.

To overcome these unphysical divergences we invoke the idea of finite-length screening. We apply the dielectric approach [32, 35, 50, 61, 62, 63, 72, 74, 75, 76], assuming infinite barriers for electrode-constituent particles [35, 50, 75, 76] and charge-carrier specular reflection at the interfaces [32, 35, 61, 62, 63, 74]. This means that (in the most general case) we have three (i = 1, 2, 3) media described by bulk dielectric functions $\epsilon_i(\mathbf{k}, \omega)$ taking into account the spatial and temporal dispersions (see Fig. 1). Here **k** is the transferred wave vector and ω is the frequency. We omit hereafter the spatial dispersion of the slab dielectric function ϵ_2 , i.e. $\epsilon_2 = \epsilon_2(\omega)$, to avoid the quantization of the quasiparticle spectrum in the interlayer. This

is true for all problems discussed below. The opposite situation for thin conducting films can be found in Refs. [63, 77].

The problem consists in calculating the image force potential energy $W_{\text{slab}}[z(t)]$ for the charge q in the slab moving normally to the interfaces, where t is the running time. In other words, the charge at any point of the trajectory z(t) induces the varying polarization charge densities on both interfaces, and interacts with them. The relevant polarization potential is $V_{\text{ind}}[x = y = 0, z(t), t]$. Thus, $W_{\text{slab}}[z(t)] = \frac{1}{2}qV_{\text{ind}}$. In the framework of the nonlocal electrostatic approach [35] using the conventional boundary conditions [73] for the electrostatic fields and inductions in the three media concerned, one obtains [78, 79]

$$W_{\text{slab}}[z(t)] = -\frac{q^2}{4\pi} \int_{-\infty}^{\infty} d\omega \frac{\exp(-i\omega t)}{\epsilon_2(\omega)}$$

$$\times \int_0^{\infty} dk_{\parallel} \frac{\exp(-2k_{\parallel}l)}{1 - \alpha_1(k_{\parallel},\omega)\alpha_3(k_{\parallel},\omega)\exp(-4k_{\parallel}l)} \int_{-\infty}^t dt' \exp(i\omega t')$$

$$\times \{\alpha_3(k_{\parallel},\omega)\exp[k_{\parallel}(z+z')] + \alpha_1(k_{\parallel},\omega)\exp[-k_{\parallel}(z+z')]$$

$$-2\alpha_1(k_{\parallel},\omega)\alpha_3(k_{\parallel},\omega)\cosh\left[k_{\parallel}(z-z')\right]\exp(-2k_{\parallel}l)\}.$$
(3)

Here k_{\parallel} is the vector **k** component along the interface, z' = z(t'), the blocks

$$\alpha_i(k_{\parallel},\omega) = \frac{\epsilon_{si}(k_{\parallel},\omega) - \epsilon_2(\omega)}{\epsilon_{si}(k_{\parallel},\omega) + \epsilon_2(\omega)} \tag{4}$$

(i = 1, 3) are expressed through the so-called surface dielectric permittivities

$$\epsilon_{si}(k_{\parallel},\omega) = \left[\frac{k_{\parallel}}{\pi} \int -\infty^{\infty} \frac{dk_z}{\mathbf{k}^2 \epsilon_i(\mathbf{k},\omega)}\right]^{-1},\tag{5}$$

and $k_z = \sqrt{\mathbf{k}^2 - k_{\parallel}^2}$. Hereafter, the arguments t in z(t) in the three-layer case or in r(t) in the two-layer one will be omitted for brevity.

It is readily seen from Eq. (3) that the response to the nonrelativistically moving charge is inertial (nonadiabatic). Actually, $W_{\rm slab}(t)$ depends on the preceding trajectory z(t' < t) due to the frequency dependences of $\alpha_i(k_{\parallel}, \omega)$. In the limiting case $\omega_s \to \infty$, when the temporal dispersion of the electrode dielectric functions is negligibly small, the expression (3) reduces to the sum with each term proportional to $\int_{-\infty}^{\infty} d\omega \exp [i\omega(t-t')] = 2\pi\delta(t-t'), \,\delta(t)$ being the Dirac delta-function. Then the image forces can be considered as static.

In this connection let us anticipate that the influence of the dielectric function frequency dependence is small. This speculation will be justified below by direct numerical calculations and the account of the plasmon impurity decay. Then any ω -dependent quantity can be expanded into the series over ω/ω_s . Since the numerical

treatment [41] shows the minor role of dissipative processes for tunneling [they are described by Im $\epsilon_i(\mathbf{k}, \omega)$] and the dielectric formalism itself is not very suitable for the consideration of such processes [59], we shall restrict ourselves to the real $\epsilon_i(\mathbf{k}, \omega)$. Thus the ω -expansion includes only even terms [80]. For present purposes it is enough to retain only the first dynamic correction.

Further simplification can be achieved for the most natural case of the vacuum or wide-gap insulating interlayer, when $\epsilon_2(\omega)$ can be approximated by the dispersionless constant. Then Eq. (3) takes the form

$$W_{\rm slab}(z) = W_{\rm slab}^{\rm st}(z) + \Delta W_{\rm slab}(z), \tag{6}$$

where

X

$$W_{\rm slab}^{\rm st}(z) = -\frac{q^2}{2\epsilon_2} \int_0^\infty dk_{\parallel} \frac{e^{-2k_{\parallel}l}}{1 - \alpha_1(k_{\parallel}, 0)\alpha_3(k_{\parallel}, 0)e^{-4k_{\parallel}l}} \left[\alpha_3(k_{\parallel}, 0)e^{2k_{\parallel}z} + \alpha_1(k_{\parallel}, 0)e^{-2k_{\parallel}z} - 2\alpha_1(k_{\parallel}, 0)\alpha_3(k_{\parallel}, 0)e^{-2k_{\parallel}l}\right]$$
(7)

is the main static image force energy term and

$$\Delta W_{\text{slab}}(z) = \frac{q^2}{4\epsilon_2} \int_0^\infty dk_{\parallel} \frac{e^{-2k_{\parallel}l}}{\left[1 - \alpha_1(k_{\parallel}, 0)\alpha_3(k_{\parallel}, 0)e^{-4k_{\parallel}l}\right]^2} \\ \times \left\{ \left[\alpha_3''(k_{\parallel}, 0) + \alpha_3^2(k_{\parallel}, 0)\alpha_1''(k_{\parallel}, 0)e^{-4k_{\parallel}l}\right] k_{\parallel} \left[\ddot{z} + k_{\parallel}\dot{z}^2\right] e^{2k_{\parallel}z} \\ - \left[\alpha_1''(k_{\parallel}, 0) + \alpha_1^2(k_{\parallel}, 0)\alpha_3''(k_{\parallel}, 0)e^{-4k_{\parallel}l}\right] k_{\parallel} \left[\ddot{z} - k_{\parallel}\dot{z}^2\right] e^{-2k_{\parallel}z} \\ - 2\left[\alpha_1''(k_{\parallel}, 0)\alpha_3(k_{\parallel}, 0) + \alpha_1(k_{\parallel}, 0)\alpha_3''(k_{\parallel}, 0)\right] k_{\parallel}^2 \dot{z}^2 e^{-2k_{\parallel}l} \right\}$$
(8)

is the dynamic correction. Here dotted and primed quantities mean time- and frequency-derivatives, respectively.

In the case $l \to \infty$ the three-layer system decomposes into a couple of two independent two-layer systems each possessing only one interface. Then for each *i*-th electrode Eqs. (6)–(8) are transformed into the following:

$$W_{\text{surf},i}(r) = W_{\text{surf},i}^{\text{st}}(r) + \Delta W_{\text{surf},i}(r), \qquad (9)$$

$$W_{\text{surf},i}^{\text{st}}(r) = -\frac{q^2}{2\epsilon_2} \int_0^\infty dk_{\parallel} \,\alpha_i(k_{\parallel},0) e^{-2k_{\parallel}r},\tag{10}$$

$$\Delta W_{\text{surf},i}(r) = \frac{q^2}{2} \int_0^\infty dk_{\parallel} k_{\parallel} \left(k_{\parallel} \dot{r}^2 - \ddot{r} \right) e^{-2k_{\parallel} r} \\ \times \left\{ \frac{\epsilon_{si}''(k_{\parallel}, 0)}{\left[\epsilon_{si}(k_{\parallel}, 0) + \epsilon_2 \right]^2} - \frac{2\epsilon_{si}'^2(k_{\parallel}, 0)}{\left[\epsilon_{si}(k_{\parallel}, 0) + \epsilon_2 \right]^3} \right\}.$$
(11)

It should be borne in mind that in actual fact the dispersionless dielectric constants differ from unity, such that our ϵ_2 are artifacts of the electrostatic approximation because for any substance it should be $\epsilon(\mathbf{k} \to \infty) \to 1$ [61, 62, 81, 82]. This conclusion stems from the fact that large k's correspond to small $|\mathbf{r} - \mathbf{r}'|$'s in the kernel $\epsilon(\mathbf{r} - \mathbf{r}')$ linking the electrostatic field and induction in the presence of the spatial dispersion [80]. Therefore, all divergences at the interfaces due to the differences between dielectric constants disappear for proper treatments, and only some smeared humps or dips of the image force energy may survive there [83]. This problem will be dealt with below in more detail for semiconducting electrodes, but hereafter in the specific calculations, to avoid further discussion of this issue, we restrict ourselves to the vacuum case $\epsilon_2 = 1$.

3. Two-layer systems

3.1 Metal-vacuum interface

Expressions (9)–(11) enable numerical calculations to be carried out for any possible metallic $\epsilon(\mathbf{k}, \omega)$, the exact form of which being unknown even in the structureless (jellium-like) case [84]. However, we confine ourselves in this section to the hydro-dynamical model of the plasma-like medium [85]

$$\epsilon(\mathbf{k},\omega) = 1 - \frac{\omega_p^2 \kappa^2}{\omega^2 \kappa^2 - \omega_p^2 \mathbf{k}^2}$$
(12)

and the uniformly accelerated motion

$$r(t) = \frac{1}{2}qFt^2\tag{13}$$

under the action of the applied electrostatic field F. Here $\omega_p = \omega_s \sqrt{2}$ is the bulk electron plasma frequency and κ is the inverse screening length. Then

$$W_{\rm surf}(\kappa r \ll 1) \approx -\frac{q^2 \kappa}{3} \left(1 - \frac{9}{16} \kappa r + \frac{3}{4} \kappa r \ln \gamma \kappa r \right) - \frac{5q^3 \kappa^2 F}{48m\omega_p^2}, \quad (14)$$

$$W_{\rm surf}(\kappa r \gg 1) \approx -\frac{q^2}{4r} \left(1 - \frac{1}{\kappa r} - \frac{qF}{m\omega_p^2 r} \right). \tag{15}$$

These dynamic corrections, calculated in the presence of the spatial dispersion, are estimated to be small [78]. Still, they lead to a substantial reduction of the field emission current density j for large (sub-threshold) F, so that the Fowler-Nordheim linear plot j/F^2 versus 1/F is violated in accordance with the experiment [86].

3.2 Semiconductor-vacuum interface

The simplest possible model for the semiconductor dielectric permittivity, taking into account the existence of the band gap and the dependence of the dielectric function

on **k** and ω , was introduced by Inkson [81]:

$$\epsilon(\mathbf{k},\omega) = 1 + \frac{\epsilon_0 - 1}{1 + (\epsilon_0 - 1)\left(\frac{\mathbf{k}^2}{\kappa^2} - \frac{\omega^2}{\omega_p^2}\right)},\tag{16}$$

where ϵ_0 is the static lattice dielectric constant, κ and ω_p are the inverse screening length and the plasma frequency of the valent electrons. This formula reproduces well the plasma-like-medium limiting cases, static (Thomas-Fermi, $\omega/k \to 0$ and $\epsilon_0 \to \infty$) and dynamic ($\omega/k \to \infty$ and $\epsilon_0 \to \infty$).

From Eqs. (9)–(11) and (16) it comes about that for $\kappa r \sqrt{\frac{\epsilon_0}{\epsilon_0 - 1}} \ll 1$

$$W_{\rm surf}(r) \approx -\frac{q^2 \kappa \sqrt{\epsilon_0(\epsilon_0 - 1)}}{4} \left\{ 1 - \frac{\epsilon_0 - 1}{\sqrt{\epsilon_0}} \tan^{-1} \frac{1}{\sqrt{\epsilon_0}} + \frac{\kappa r}{\sqrt{\epsilon_0(\epsilon_0 - 1)}} \ln(\gamma \kappa r \sqrt{\frac{\epsilon_0}{\epsilon_0 - 1}}) + \frac{\kappa r}{2\sqrt{\epsilon_0(\epsilon_0 - 1)}} \right\}$$
$$\times [(\epsilon_0 - 2) - (\epsilon_0^2 - 1) \ln \frac{\epsilon_0 + 1}{\epsilon_0}] \left\{ -\frac{q^3 \kappa^2 F \epsilon_0^2}{16m \omega_p^2} \left[\frac{3}{\epsilon_0} + 2(\epsilon_0 - 1) \ln \frac{\epsilon_0 + 1}{\epsilon_0} - 2 \right] \right\}$$
(17)

and for $\kappa r \sqrt{\frac{\epsilon_0}{\epsilon_0-1}} \gg 1$

$$W_{\rm surf}(r) \approx -\frac{q^2(\epsilon_0 - 1)}{4r(\epsilon_0 + 1)} \left[1 - \frac{\sqrt{\epsilon_0(\epsilon_0 - 1)}}{\kappa r(\epsilon_0 + 1)} - \frac{qF(\epsilon_0 - 1)}{m\omega_p^2 r(\epsilon_0 + 1)} \right].$$
 (18)

The asymptotics (14)–(15) stem from Eqs. (17)–(18) in the limit $\epsilon_0 \rightarrow \infty$, *i.e.*, for the infinite ionicity formally appropriate to a metal. One can see that the conventional description of the image forces near the semiconductor surface [29, 32, 38, 39, 87] is recovered only at large distances. The dynamic corrections are of the same type as for the metal with itinerant electrons, although in the present case all the electrons are bound. This similarity resembles one for the electron plasma response in the energy-loss experiments for both metals and small-gap semiconductors [30, 32, 38, 87, 88, 89].

4. Three-layer systems

4.1 Metal-vacuum-metal structures

The calculations in the general case of a three-layer sandwich with metallic covers (M-I-M) can be carried out in the same manner as for the single interface, starting

Tunneling in Two- and Three-layer Systems ...



Fig. 2. The image force potential barriers for the junctions Sb-vacuum–Sb with and without dynamic corrections. $F/F_0 = 0.05$ and $\delta b = 0.5, 1$, and 2 (curves 1–3, respectively). See notations in the text.

from the hydrodynamic approximation (12) and Eqs. (7) and (8). For thin symmetrical sandwiches, *i.e.*, for $\kappa_1 = \kappa_3 = \kappa$, $\omega_{p1} = \omega_{p2} = \omega_p$, and $\delta = \kappa l \ll 1$, it follows

$$W_{\rm slab}^{\rm st}(\xi) \approx \frac{1}{2} q^2 \kappa \{1 + \delta$$
(19)
 $\times [\ln(2\delta\gamma) - 1 + \frac{1}{2}(1-\xi)\ln(1-\xi) + \frac{1}{2}(1+\xi)\ln(1+\xi)]\},$
$$\Delta W_{\rm slab}(\xi) \approx \frac{q^3 F \kappa^2 \delta}{12m\omega_n^2} (4 + 11\xi).$$
(20)

Here $\xi = z/l$, and hence $-1 \le \xi \le 1$, and the charge moves in the vacuum gap from the left to the right electrode forced by the applied electrostatic field F. The main static term is symmetrical about ξ , and for $\delta \to 0$ tends to $-q^2\kappa/2$ which is exactly the inner electrostatic potential energy in the Thomas-Fermi approximation [52]. This is the charge energy averaged over the crystal volume and reckoned from the vacuum level [90].

The dynamic correction is *asymmetrical* with respect to the origin and alternating. The asymmetry is associated with the accelerated character of the motion. In particular, for uniform motion [35] Eq. (8) causes a symmetrical and positive dynamic correction that suppresses the static image forces for all ξ . Our calculations show [74] that in M–I–M structures the dynamic corrections are small for good metals. On the contrary, for semimetallic Sb electrodes with $\omega_p \approx 4.15 \cdot 10^{14} \text{ s}^{-1}$ and $\kappa \approx 1.66 \cdot 10^6 \text{ cm}^{-1}$ the deviations from the static behavior may be conspicuous (see Fig. 2). Here F = V/2l is measured in the units of $F_0 = 10^8 \text{ V cm}^{-1}$ typical for the given problem, V is the potential interelectrode difference.

4.2 Semiconductor-vacuum-semiconductor structures

On the basis of Eqs. (7) and (8) it is possible to obtain the dynamic image force energy profiles in the slab between nondegenerate semiconductors with the dielectric permittivity (16). For thin symmetrical sandwiches with $\delta \sqrt{\frac{\epsilon_0}{\epsilon_0 - 1}} \ll 1$ the corresponding expressions take the form

$$W_{\rm slab}^{\rm st}(\xi) \approx$$

$$-\frac{q^{2}\kappa}{2}\left\{\sqrt{\frac{\epsilon_{0}-1}{\epsilon_{0}}}+\delta\left[\ln\left(\gamma\delta\sqrt{\frac{\epsilon_{0}}{\epsilon_{0}-1}}\right)+\frac{\epsilon_{0}-1}{\epsilon_{0}}\left(\ln\ 2-\frac{1}{2}\right)-\frac{1}{2}\right.\\\left.+\frac{1}{2}(1-\xi)\ln(1-\xi)+\frac{1}{2}(1+\xi)\ln(1+\xi)\right]\right\},$$
(21)

$$\Delta W_{\rm slab}(\xi) \approx \delta \frac{q^3 F \kappa^2}{12m\omega_p^2} \sqrt{\frac{\epsilon_0 - 1}{\epsilon_0}} \{ 4 + \frac{\xi}{(\epsilon_0 - 2)^3} [11\epsilon_0^3 - 34\epsilon_0^2 + 63\epsilon_0 - 48 + \frac{6\epsilon_0(\epsilon_0 - 1)(4\epsilon_0 - 3)}{\epsilon_0 - 2} f(\epsilon_0)] \},$$
(22)

where

$$f(\epsilon_0) = \begin{cases} \sqrt{\frac{2-\epsilon_0}{\epsilon_0}} \tan^{-1}\sqrt{\frac{2-\epsilon_0}{\epsilon_0}} & \text{(for } 1 < \epsilon_0 \le 2), \\ -\frac{1}{2}\sqrt{\frac{\epsilon_0-2}{\epsilon_0}} \ln \frac{\sqrt{\epsilon_0}+\sqrt{\epsilon_0-2}}{\sqrt{\epsilon_0}-\sqrt{\epsilon_0-2}} & \text{(for } \epsilon_0 > 2). \end{cases}$$
(23)

These cumbersome expressions reduce to those for the metallic covers [Eqs. (19) and (20)] in the previously described limit $\epsilon_0 \to \infty$. The value 2 for ϵ_0 is *not* the singular point, contrary to what might be expected, the quantity in brackets being quite smooth. On the contrary, it ranges from 11 for $\epsilon_0 \to \infty$ to 8 for $\epsilon_0 \to 1$. Once again, it should be noted that both static image forces and dynamic corrections are very similar for metallic and semiconducting heterostructures.

5. Justification of the adopted approach

The dynamic corrections appeared to be substantial but small enough to justify our perturbation approach, described in Section 2. and in more detail in Refs. [78, 79]. However, some doubts may remain concerning the applicability of the perturbation procedure for the image forces in the case of the emitted projectiles when the real surface plasmon avalanche is left in the wake, according to the semiclassical theory [35, 66, 91]. Then the image force energy includes long-range oscillating terms. This behavior is conserved in the quantum-mechanical theory when the recoil effect is small (the "above-threshold" situation) [59, 60, 71, 72]. Such a treatment leads to a spatial decay proportional to $\exp\left(-r\sqrt{\frac{2m\omega_s}{\hbar}}\right)$ [71], where \hbar is the Planck's constant. This decay is weak for electrodes with small current-carrier densities. Nevertheless, there is an important factor, namely, the collision plasmon damping, which

results in the restoration of the conventional image force energy power-law dependence in the asymptotic region [92].

Let us consider the uniform quasiclassical motion of a charge near the metal– vacuum interface [35] with the dielectric permittivity obtained in the framework of the kinetic equation [93]

$$\epsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i\nu)}.$$
(24)

Here ν is the inverse relaxation time, and the spatial dispersion is neglected as being insignificant in this case. Incidentally, the problem of the rigorous introduction of the damping factor into the dielectric permittivity of the medium with temporal and spatial dispersions is both far from being solved and far from being unambiguously formulated [94, 95].

The starting expression for the image force energy for a charge q moving in vacuum normally to the metal surface in accordance with the law r(t) has the form (compare with Ref. [35])

$$W[r(t)] = -\frac{q^2 \tilde{\omega}_s}{2} \int_{-\infty}^t dt' \, \frac{\sin \, \tilde{\omega}_s(t-t')}{r(t)+r(t')} \, \exp\left[-\frac{\nu}{2}(t-t')\right], \tag{25}$$

where

$$\tilde{\omega}_s = \sqrt{\frac{\omega_p^2}{2} - \frac{\nu^2}{4}}.$$
(26)

For the constant speed v it is convenient to measure distances in units of $L = 2\pi v / \tilde{\omega}_s$. Then the problem includes a single input dissipation parameter $\beta = \nu / \omega_p$. The most "dangerous" set-up is the motion of the emitted particle. Consequently, elementary calculations lead to

$$W_{\rm emit}(r \to 0) \approx -\frac{q^2 \tilde{\omega}_s^2 r}{2v^2} \left[2 \ln 2 - 1 - \frac{\nu r}{v} \left(2 \ln 2 - \frac{5}{4} \right) \right], \qquad (27)$$

$$W_{\text{emit}}(r \to \infty) \approx$$

$$-\frac{q^2}{4r} \{1 - \exp(-\frac{\nu r}{2v}) [2\cos\frac{\tilde{\omega}_s r}{v} + \frac{\nu}{\tilde{\omega}_s}(1 - \frac{2v}{\nu r})\sin\frac{\tilde{\omega}_s r}{v}]\}.$$

$$(28)$$

One sees that the saturated value of $W_{\rm emit}(0)$ is almost unaltered by the plasmon damping, whereas the huge oscillations, totally distorting the image forces for $r \rightarrow \infty$ in the situation $\nu = 0$ [35], are rapidly damped in the real case $\nu \neq 0$. Asymptotics (27) and (29) are complemented by numerical calculations shown in Fig. 3. It is remarkable that the damping factor does not depend on ω_s and becomes especially large for slow projectiles, *e.g.*, protons [32, 88].

The stabilizing role of the plasmon finite lifetime for image forces should manifest itself also for a slab geometry and other laws of charge motion.



Fig. 3. Dependences of the dimensionless image force energy $w = 2vW/q^2\tilde{\omega}_s$ on the dimensionless distance ξ for the particle emitted from the metal with different plasmon dissipative parameters β . See notations in the text.

6. Tunnel currents in three-layer systems

The energy level diagram for the electrically biased tunnel junction taking into account the dynamic image forces is shown in Fig. 4. Bearing in mind the actual



Fig. 4. Schematic tunnel barriers in a thin symmetrical M–I–M biased junction taking into account static and dynamic image forces. See notations in the text.

smallness of the dynamic corrections, we may carry out the whole analysis in the traditional manner [96], *i.e.*, neglecting the temperature dependence of the Fermi-Dirac distribution function and employing the semiclassical approximation and the saddle point method. Then the electron tunnel current density j through a symmetric M–I–M junction can be written in the form

$$j = j_0(\eta) - j_0(\eta - eV),$$
 (29)

Tunneling in Two- and Three-layer Systems ...

$$j_0(\eta) = \frac{em}{2\pi^2\hbar^3} \left(\frac{\partial I}{\partial E_z}\right)_{E_z=\eta}^{-2} \exp[-I(\eta)].$$
(30)

Here e is the elementary charge, $I(E_z)$ is the tunneling exponent which determines the JWKB tunneling rate $D(E_z) = \exp[-I(E_z)]$, E_z is the energy of the electron motion across the junction, and η is the Fermi energy of the metal measured from its conduction band bottom.

In the semiclassical approximation adopted here, the exponent $I(E_z)$ is determined by the co-ordinate dependence of the velocity $v(z) = |\dot{z}|$ for the sub-barrier motion [1, 96]:

$$I(E_z) = \frac{2m}{\hbar} \int_{z_1}^{z_2} dz \, v(z),$$
(31)

$$\frac{1}{2}mv^2(z) = U(z) - E_z,$$
(32)

where U(z) is the potential electron energy in the interlayer, and $z_{1,2}$ are the turning points at which $v(z_{1,2}) = 0$. The tunneling time enters into consideration implicitly through Eq. (31) because the corresponding semiclassical tunneling time is $\tau_{\text{tun}}^{\text{semicl}} = \int_{z_1}^{z_2} v^{-1}(z) dz$ [97]. $\tau_{\text{tun}}^{\text{semicl}}$ is the time that a particle with a real velocity v(z)would take to traverse the barrier [98]. Further subtleties, *e.g.*, concerning different tunneling times [1, 2, 3, 4, 5, 8, 11] do not interfere, because the dynamic corrections are small. Rewriting Eq. (8) for the dynamic corrections in the following manner

$$\Delta W(z) = \frac{1}{2}m\dot{z}^{2}\rho_{1}(z) + m\ddot{z}z\rho_{2}(z), \qquad (33)$$

i.e., introducing the functions $\rho_1(z)$ and $\rho_2(z)$, the potential energy U(z) reads

$$U(z) = \eta + \mu - eF(z+l) + W^{\rm st}(z) + \frac{1}{2}mv^2\rho_1(z) + m\dot{v}z\rho_2(z), \qquad (34)$$

where μ is the work function, $W^{\text{st}}(z)$ is obtained from Eq. (7) with $\alpha_1(k_{\parallel}, 0) = \alpha_3(k_{\parallel}, 0)$, and $\rho_1(z)$ and $\rho_2(z)$ are even functions of z in the case of identical electrodes. The value of v^2 can be immediately found from Eqs. (31) and (34). Differentiating of the kinetic energy $\frac{1}{2}mv^2$ with respect to time, and making allowance for the small dynamic correction, leads to [79]

$$m\dot{v} \approx -eF,$$
 (35)

i.e., the sub-barrier "motion" is decelerated by the electric field.

It should be stressed that this treatment of the electron in the classically forbidden area as moving is not more bizarre or inconsistent than its "adiabatically immobile" version in the conventional semiclassical theory involving static image forces [39, 86, 96]. The classical motion studied by us is analogous to that in the inverted effective potential of the path-integral approach [44]. The introduction of the electrostatic field as a source of motion is our new key point [37, 78, 79] overlooked in other investigations of the electron tunneling [40, 41, 42, 44, 56, 67, 70]. The only attempt [69] known to us to make allowance for the direct field influence on the image forces contains only a guess not brought to completion.

Solving the system of Eqs. (29)–(35) for thin barriers $\kappa l \ll 1$, when the functions $\rho_1(z)$ and $\rho_2(z)$ are independent of z, it is possible to obtain a formula similar to the classical Fowler-Nordheim formula [79]

$$j(F) = \frac{e^3 F^{*2}}{(4\pi)^2 \hbar \mu^*} \exp\left(-\frac{4\sqrt{2m\mu^{*3}}}{3\hbar e F^*}\right),$$
(36)

but with renormalized values of effective work function

$$\mu^* = \frac{\mu + \tilde{W}^{\text{st}} + eFl\rho_2}{1 - \rho_1} \tag{37}$$

and external field

$$F^* = \frac{1+\rho_2}{1-\rho_1}F.$$
(38)

 \tilde{W}^{st} is the average of W^{st} across the junction

$$\tilde{W}^{\rm st} = \frac{1}{2l} \int_{-l}^{l} dz \, W^{\rm st}(z).$$
(39)

Eq. (36) was obtained for strong electric fields, when $2eF^*l > \mu^*$. The opposite case of small voltages ($2eF^*l \ll \mu^*$) can be found elsewhere [79].

The diagram in Fig. 4 shows that the dynamic corrections increase the height and width of the tunnel barrier formed by the applied electrostatic field and static image forces. Therefore, the corresponding tunnel current is reduced. The field dependence of μ^* leads to deviations from the linear Fowler-Nordheim plot $\ln (j/F^2) \sim F^{-1}$ toward smaller current values. Similar deviations, observed for the cold emission from a metal to the vacuum [86], were explained in the same manner [78].

7. Conclusions

The main conclusion consists in the important role of the dynamic character of the image forces in tunneling both for two- and three-layer systems. The nonadiabaticity is due to the close orders of the tunneling time and the inverse frequency of surface plasmons excited in electrodes by charges moving in external electric fields. Although small, the dynamic corrections are responsible for deviations from the Fowler-Nordheim law in cold emission, the latter being among the early manifestations of the tunneling phenomenon itself.

This work was supported in part by the Ukrainian State Foundation for Fundamental Research (grant 2.4/100).

References

- Tunneling and Its Implications, edited by D. Mugnai, A. Ranfagni, and L. S. Schulman (World Scientific, Singapore, 1997).
- [2] E. H. Hauge and J. A. Støvneng, Rev. Mod. Phys. 61, 917 (1989).
- [3] M. Büttiker, in *Electronic Properties of Multilayers and Low-Dimensional Semiconductors Structures*, edited by J. M. Chamberlain, L. Eaves, and J. C. Portal (Plenum Press, New York, 1990), p. 297.
- [4] V. S. Olkhovsky and E. Recami, Phys. Rep. 214, 339 (1992).
- [5] R. Landauer and Th. Martin, Rev. Mod. Phys. 66, 217 (1994).
- [6] G. Nimtz and W. Heitmann, Progr. Quant. Electr. 21, 81 (1997).
- [7] R. Y. Chiao and A. M. Steinberg, in *Progress in Optics XXXVII*, edited by E. Wolf (Elsevier, Amsterdam, 1997), p. 345.
- [8] J. Jakiel, V. S. Olkhovsky, and E. Recami, Phys. Lett. A 248, 156 (1998).
- [9] H. A. Fertig, Phys. Rev. Lett. 65, 2321 (1990).
- [10] M. Ya. Azhel, Solid State Commun. 91, 439 (1994).
- [11] T. Figielski, Solid State Commun. 94, 113 (1995).
- [12] E. U. Condon, Rev. Mod. Phys. 3, 43 (1931).
- [13] L. A. MacColl, Phys. Rev. 40, 621 (1932).
- [14] R. H. Fowler and L. W. Nordheim, Proc. Roy. Soc. A 119, 173 (1928).
- [15] L. M. Baskin, R. Z. Bakhtizin, V. G. Valeev, and G. N. Fursei, in *Field, Thermionic and Secondary Electron Emission Spectroscopy*, edited by G. N. Fursei (Nauka, Moscow, 1990), p. 293, in Russian.
- [16] S. A. Mikhailov and V. A. Volkov, Pis'ma Zh. Eksp. Teor. Fiz. 61, 508 (1995).
- [17] J. R. Oppenheimer, Phys. Rev. 31, 66 (1928).
- [18] K. I. Zamaraev, R. F. Khairutdinov, and V. P. Zhdanov, *Electron Tunneling in Chemistry* (Nauka, Novosibirsk, 1985), in Russian.
- [19] G. Gamow, Z. Phys. 51, 204 (1928).
- [20] R. W. Gurney and E. U. Condon, Nature 122, 439 (1928).
- [21] S. Bjørnholm and J. E. Lynn, Rev. Mod. Phys. 52, 725 (1980).
- [22] N. Takigawa and K. Hagino, Progr. Theor. Phys. 101 (1996).
- [23] A. B. Balantekin and N. Takigawa, Rev. Mod. Phys. 70, 77 (1998).
- [24] V. I. Goldanskii, L. I. Trakhtenberg, and V. N. Flyorov, *Tunnel Phenomena in Chemical Physics* (Nauka, Moscow, 1986), in Russian.
- [25] V. A. Benderskii, V. I. Goldanskii, and D. E. Makarov, Phys. Rep. 233, 195 (1993).
- [26] D. G. Bourgin, Proc. Nat. Acad. Sci. USA 15, 357 (1929).
- [27] S. Z. Roginsky and L. Rosenkevich, Z. Phys. Chem. B 10, 47 (1930).
- [28] A. A. Lucas and M. Šunjić, Progr. Surf. Sci. 2, 76 (1972).
- [29] E. Evans and D. L. Mills, Phys. Rev. B 8, 4004 (1973).
- [30] Ju. I. Juaristi, F. J. García de Abajo, and P. M. Echenique, Phys. Rev. B 53, 13839 (1996).
- [31] H. Winter, J. Phys.: Condens. Matter 8, 10149 (1996).
- [32] A. Arnau, F. Aumayr, P. M. Echenique, M. Grether, W. Heiland, J. Limburg, R. Morgenstern, P. Roncin, S. Schippers, R. Schuch, N. Stolterfoht, P. Varga, T. J. M. Zouros, and H. P. Winter, *Surf. Sci. Rep.* 27, 113 (1997).
- [33] C. Denton, J. L. Gervasoni, R. O. Barrachina, and N. R. Arista, Phys. Rev. A 57, 4498 (1998).
- [34] V. U. Nazarov and Yu. V. Luniakov, Phys. Rev. B 52, 12414 (1995).
- [35] J. Heinrichs, Phys. Rev. B 8, 1346 (1973).
- [36] Yu. A. Romanov and V. Ya. Alyoshkin, Fiz. Tverd. Tela 23, 86 (1981).
- [37] A. M. Gabovich, Fiz. Tverd. Tela 25, 1885 (1983).
- [38] Ph. Lambin, A. A. Lucas, and J. P. Vigneron, Surf. Sci. 182, 567 (1987).
- [39] P. J. Jennings and R. O. Jones, Adv. Phys. 37, 341 (1988).
- [40] B. N. J. Persson and A. Baratoff, Phys. Rev. B 38, 9616 (1988).
- [41] B. G. R. Rudberg and M. Jonson, Phys. Rev. B 43, 9358 (1991).
- [42] M. Šunjić and L. Marušić, Phys. Rev. B 44, 9092 (1991).

- [43] K. L. Sebastian and G. Doyen, Phys. Rev. B 47, 7634 (1993).
- [44] N. Klipa and M. Šunjić, Phys. Rev. B 52, 12408 (1995).
- [45] D. Šestović, L. Marušić, and M. Šunjić, Phys. Rev. B 55, 1741 (1997).
- [46] M. Sumetskii, Solid State Commun. 74, 877 (1990).
- [47] D. V. Averin, Phys. Rev. B 50, 8934 (1994).
- [48] F. Sols and R. H. Ritchie, Phys. Rev. B 35, 9314 (1987).
- [49] D. B. Tran Thoai, E. Zeitler, and D. Nam, Phys. Status Solidi B 158, 557 (1990).
- [50] A. A. Kornyshev, A. I. Rubinshtein, and M. A. Vorotyntsev, J. Phys. C 11, 3307 (1978).
- [51] P. R. Rimbey, Surf. Sci. 85, 166 (1979).
- [52] A. V. Sidyakin, Zh. Eksp. Teor. Fiz. 58, 573 (1970).
- [53] D. B. Tran Thoai and E. Zeitler, Phys. Status Solidi B 146, 137 (1988).
- [54] J. Mahanty and P. Summerside, J. Phys. F 10, 1013 (1980).
- [55] M. T. Michalewicz and J. Mahanty, Phys. Lett. A 116, 392 (1986).
- [56] P. M. Echenique, A. Gras-Marti, J. R. Manson, and R. H. Ritchie, Phys. Rev. B 35, 7357 (1987).
- [57] J. Mahanty, K. N. Pathak, and V. V. Paranjape, Solid State Commun. 54, 649 (1985).
- [58] J. Mahanty, K. N. Pathak, and V. V. Paranjape, Phys. Rev. B 33, 2333 (1986).
- [59] J. R. Manson and R. H. Ritchie, Phys. Rev. B 24, 4867 (1981).
- [60] F. Sols and R. H. Ritchie, Surf. Sci. 194, 275 (1988).
- [61] A. I. Voitenko, A. M. Gabovich, and L. G. Il'chenko, Fiz. Tverd. Tela 23, 1531 (1981).
- [62] A. M. Gabovich and A. I. Voitenko, Phys. Status Solidi B 100, 407 (1982).
- [63] A. M. Gabovich, L. G. Il'chenko, E. A. Pashitskii, and Yu. A. Romanov, Surf. Sci. 94, 179 (1980).
- [64] M. Šunjić, G. Toulouse, and A. A. Lucas, Solid State Commun. 11, 1629 (1972).
- [65] C. Shah and G. Mukhopadhyay, Solid State Commun. 48, 1035 (1983).
- [66] R. Ray and G. D. Mahan, Phys. Lett. A 42, 301 (1972).
- [67] M. Jonson, Solid State Commun. 33, 743 (1980).
- [68] P. R. Rao, C. Shah, and G. Mukhopadhyay, Pramāna 23, 119 (1984).
- [69] R. A. Young, Solid State Commun. 45, 263 (1983).
- [70] J.-W. Wu and G. D. Mahan, Phys. Rev. B 28, 4839 (1983).
- [71] X.-Y. Zheng, R. H. Ritchie, and J. R. Manson, *Phys. Rev. B* 39, 13510 (1989).
- [72] P. M. Echenique, R. H. Ritchie, N. Barberan, and J. Inkson, Phys. Rev. B 23, 6486 (1981).
- [73] W. R. Smythe, Static and Dynamic Electricity (McGraw-Hill, New York, 1950).
- [74] A. M. Gabovich, L. G. Il'chenko, and E. A. Pashitskii, Fiz. Tverd. Tela 21, 1683 (1979).
- [75] D. M. Newns, *Phys. Rev. B* 1, 3304 (1970).
- [76] V. U. Nazarov, Phys. Rev. B 49, 10663 (1994).
- [77] L. G. Il'chenko, E. A. Pashitskii, and Yu. A. Romanov, Fiz. Tverd. Tela 22, 2700 (1980).
- [78] A. M. Gabovich, V. M. Rozenbaum, and A. I. Voitenko, Surf. Sci. 186, 523 (1987).
- [79] A. I. Voitenko, A. M. Gabovich, and V. M. Rozenbaum, Fiz. Nizk. Temp. 22, 86 (1996).
- [80] L. D. Landau and E. M. Lifshits, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1982), in Russian
- [81] J. C. Inkson, J. Phys. C 4, 591 (1971).
- [82] A. A. Kornyshev, Electrochim. Acta 26, 1 (1981).
- [83] A. M. Gabovich and A. I. Voitenko, *Electrochim. Acta* 31, 777 (1986).
- [84] V. D. Gorobchenko and E. G. Maksimov, Usp. Fiz. Nauk 130, 65 (1980).
- [85] G. Barton, Rep. Prog. Phys. 42, 963 (1979).
- [86] R. H. Good, Jr and E. W. Müller, in *Encyclopedia of Physics, Vol. 21*, edited by S. Flügge (Springer Verlag, Berlin, 1956), p. 176.
- [87] L. D. Marks, Solid State Commun. 43, 727 (1982).
- [88] A. Gras-Marti, P. M. Echenique, and R. H. Ritchie, Surf. Sci. 173, 310 (1986).
- [89] J.-P. Briand, S. Thuriez, G. Giardino, G. Borsoni, V. Le Roux, M. Froment, M. Eddrief, C. de Villeneuve, B. D'Etat-Ban, and C. Sébenne, *Phys. Rev. A* 55, 2523 (1997).
- [90] A. Sánchez and M. A. Ochando, J. Phys. C 18, 33 (1985).
- [91] P. M. Echenique and J. B. Pendry, J. Phys. C 8, 2936 (1975).

- [92] A. M. Gabovich, V. M. Rozenbaum, and A. I. Voitenko, Phys. Status Solidi B 214, 29 (1999).
- [93] Yu., L. Klimontovich, Statistical Physics (Nauka, Moscow, 1982), in Russian.
- [94] N. D. Mermin, Phys. Rev. B 1, 2362 (1970).
- [95] A. K. Das, J. Phys. F 5, 2035 (1975).
- [96] E. L. Wolf, Principles of Electron Tunneling Spectroscopy (Oxford University Press, New York, 1985).
- [97] M. Büttiker and R. Landauer, Phys. Rev. Lett. 49, 1739 (1982).
- [98] M. Büttiker, Phys. Rev. B 27, 6178 (1983).

Tunneling in the Wigner Representation

Bilha Segev Department of Chemistry, Ben-Gurion University of the Negev P.O.B. 653, Beer-Sheva 84105, Israel

The Wigner function is used to study the purely quantum time evolution of wave packets. Wave packets incident on potential barriers or undergoing quantum transitions between energy surfaces are studied, demonstrating in both cases the utility of the Wigner representation for describing pure quantum effects with no classical counterparts.

Keywords: tunneling, phase-space, causality, Wigner function, time

1. Introduction

In 1932 Wigner wrote a paper entitled "On the quantum corrections for thermodynamics equilibrium" in which he introduced what later became known as the Wigner function. In [1] Wigner writes:

"If a wave function $\Psi(x_1, ..., x_n)$ is given one may build the following expression

$$\mathcal{P}(x_1, ..., x_n; p_1, ..., p_n) = \frac{(\frac{1}{h\pi})^n \int_{-\infty}^{\infty} ... \int dy_1 ... dy_n \Psi(x_1 + y_1, ..., x_n + y_n)^*}{\times \Psi(x_1 - y_1, ..., x_n - y_n) \ e^{2i(p_1y_1 + ... + p_ny_n)/h}}$$
(1)

and call it the probability-function of the simultaneous values of $x_1, ..., x_n$ for the coordinates and $p_1, ..., p_n$ for the momenta". "Of course $\mathcal{P}(x_1, ..., x_n; p_1, ..., p_n)$ cannot be really interpreted as the simultaneous probability for coordinates and momenta, as is clear from the fact, that it may take negative values. But of course this must not hinder its use in calculations as an auxiliary function which obeys many relations we would expect from such a probability."

Since then, the Wigner function has been used for various applications, and many papers and reviews have been written about it, [2]-[12]. Traditionally, the similarity to classical distributions has encouraged applications in semiclassical theories. Here I take a complementary viewpoint, and review recent work by M.S. Marinov and myself [13]-[14], as well as by E.J. Heller and myself [15] emphasizing the application of the Wigner representation to purely quantum effects with no classical counterpart. This work includes applications to scattering, wave-packet propagation, and tunneling time, and to energy transfer processes within a single molecule, includ-

B. Segev

ing, in particular, nonclassical Frank Condon factors and radiationless transitions in polyatomic molecules. What these different phenomena have in common is their nonclassical nature, which is treated here within a phase-space Wigner approach to tunneling.

2. The Wigner representation

In the Wigner Representation a quantum state given by the density matrix $\hat{\rho}_t$ is represented by a phase-space quasi-distribution. (I use units with $\hbar = 1$).

$$\rho_t(q,p) = \int d\eta \, \langle q + \frac{\eta}{2} | \hat{\rho}_t | q - \frac{\eta}{2} \rangle e^{-ip\eta} \,. \tag{2}$$

All integrals are from $\pm \infty$. An operator \hat{A} is represented by its Weyl transform:

$$A(q,p) = \int d\eta \, \langle q + \frac{\eta}{2} | \hat{A} | q - \frac{\eta}{2} \rangle e^{-ip\eta} \,. \tag{3}$$

Expectation values are given by integration:

$$\langle \hat{A} \rangle_t = \frac{1}{2\pi} \int \int dq dp \ \rho_t(q, p) A(q, p) \ , \tag{4}$$

and projection gives the probabilities in coordinate and momentum space:

$$\mathcal{P}_t(q) = \frac{1}{2\pi} \int dp \, \rho_t(q, p) \quad \left(= |\Psi(q)|^2\right) ,$$
 (5)

$$\mathcal{P}_t(p) = \int dq \,\rho_t(q,p) \quad \left(=|\tilde{\Psi}(p)|^2\right) \,. \tag{6}$$

The expression for pure states is given in parentheses, but the discussion is not limited to pure states.

3. Dynamics in the Wigner representation

Time evolution in quantum mechanics is given by the time evolution operator:

$$\hat{\rho}_t = \hat{U}(t)\hat{\rho}_0 \hat{U}^{\dagger}(t) , \ \hat{U}(t) \equiv \exp(-i\hat{H}t) ,$$
(7)

with the Hamiltonian \hat{H} . In the Wigner representation, this time evolution is given by the phase-space propagator defined and applied in the following way:

$$\rho_t(q,p) = \frac{1}{2\pi} \int \int dq_0 dp_0 \mathcal{L}_t(q,p;q_0,p_0) \rho_0(q_0,p_0) \quad . \tag{8}$$

The propagators are integrable and normalized:

$$\frac{1}{2\pi} \int dq \int dp \mathcal{L}_t^{\zeta} = \frac{1}{2\pi} \int dq_0 \int dp_0 \mathcal{L}_t^{\zeta} = 1 , \qquad (9)$$

and are bilinear transforms of matrix elements of the evolution operator:

$$\mathcal{L}_{t}^{W}(q, p; q_{0}, p_{0}) = \frac{1}{2\pi} \int dq \int dp' \exp[i(q'p + q_{0}p')] \\ \times \left\langle q - \frac{1}{2}q' \left| \hat{U}(t) \right| p_{0} - \frac{1}{2}p' \right\rangle \left\langle p_{0} + \frac{1}{2}p' \left| \hat{U}^{\dagger}(t) \right| q + \frac{1}{2}q' \right\rangle .$$
(10)

The transition probability between an initial state $\hat{\rho}_i$ at t = 0 and a final state $\hat{\rho}_f$ at a later time t is:

$$w_{if}(t) = \left(\frac{1}{2\pi}\right)^2 \int dq dp \int dq_0 dp_0 \ \rho_f(q, p) \mathcal{L}_t(q, p; q_0, p_0) \rho_i(q_0, p_0)$$

= $Tr[\hat{\rho}_f \hat{U}(t) \hat{\rho}_i U^+(t)],$ (11)

while the instantaneous transition probability is:

$$w_{if}^{0} = Tr[\hat{\rho}_{f}\hat{\rho}_{i}] = \frac{1}{2\pi} \int dq dp \ \rho_{f}(q,p)\rho_{i}(q,p) \,. \tag{12}$$

The propagator in phase space is analogous to the Dirac propagator in coordinate space, $\langle q | \hat{U} | q_0 \rangle$, and a similar propagator in momentum space, $\langle p | \hat{U} | p_0 \rangle$. Examples of different evolution kernels include the propagators for free motion:

$$\langle q|\hat{U}|q_0\rangle = \sqrt{\frac{m}{2\pi i t}} \exp\left[\frac{im}{2t}(q-q_0)^2\right],$$
 (13)

$$\langle p|\hat{U}|p_0\rangle = e^{-i\frac{p^2}{2m}t}\delta(p-p_0), \qquad (14)$$

$$\mathcal{L}_{t}^{f}(q, p; q_{0}, p_{0}) = \delta(p - p_{0})\delta(q - tp/m - q_{0}), \qquad (15)$$

$$\rho_t^f(q,p) = \rho_0(q - tp/m, p), \qquad (16)$$

and the propagator of the harmonic oscillator with unit mass, m = 1, and unit frequency, $\omega = 1$:

$$\langle q|\hat{U}|q_0\rangle = (2\pi i \sin t)^{-\frac{1}{2}} \exp\left[i\frac{(q^2+q_0^2)\cos t-2qq_0}{2\sin t}\right],$$
 (17)

$$\langle p|\hat{U}|p_0\rangle = (2\pi i \sin t)^{-\frac{1}{2}} \exp\left[i\frac{(p^2+p_0^2)\cos t-2pp_0}{2\sin t}\right],$$
 (18)

$$\mathcal{L}_t(q, p; q_0, p_0) = \tag{19}$$

 $2\pi\delta(p-p_0\cos t+q_0\sin t)\delta(q-q_0\cos t-p_0\sin t).$

In the cases of both free propagation and the harmonic oscillator, the dynamics in phase space is extremely simple. The propagator $L_t(q, p; q_0, p_0)$ is a δ function defining a one-to-one correspondence between initial and final phase-space points. The Wigner functions propagate in both cases in a completely classical manner, and each point of the Wigner function propagates on a classical trajectory. This, however, is not the generic case. Two counter examples include: tunneling through a δ potential barrier and tunneling through a modified Pöschl-Teller barrier. The propagator for the narrow potential barrier, $V^{\delta}(q) = v_0 \delta(q)$, is:

$$\mathcal{L}_{t} = \delta(p - p_{0})\delta\left(q_{0} + t\frac{p}{m} - q\right)$$

$$- \delta(p - p_{0})\theta\left(q_{0} + t\frac{p}{m} - q\right)\left[2v_{0}\sqrt{1 + \left(\frac{v_{0}}{4p}\right)^{2}}\right]$$

$$\times \exp\left[-v_{0}\left(q_{0} + t\frac{p}{m} - q\right)\right]$$

$$\times \cos\left[2p\left(q_{0} + t\frac{p}{m} - q\right) - \arctan\left(\frac{v_{0}}{4p}\right)\right]$$

$$+ \delta(p + p_{0})\theta\left(q_{0} + t\frac{p}{m} + q\right)\left[\frac{v_{0}^{2}}{2p}\right]$$

$$\times \exp\left[-v_{0}\left(q_{0} + t\frac{p}{m} + q\right)\right]\sin\left[2p\left(q_{0} + t\frac{p}{m} + q\right)\right]. \quad (20)$$

The propagator for a general one-dimensional potential barrier is:

$$\mathcal{L}_t(q, p; q_0, p_0) = \mathcal{T} + \mathcal{R} + \mathcal{S} , \qquad (21)$$

where for the modified Pöschl-Teller barrier, $V^{\rm PT}(q) = V_0^2 / \cosh^2(q/s)$:

$$\mathcal{T} = \delta(p - p_0)\delta\left(q - q_0 - t\frac{p}{m}\right) - \delta(p - p_0)\theta\left(q - q_0 - t\frac{p}{m}\right)\mathcal{F}_t(\nu,\omega) .$$
(22)

 $\nu = 2p_0 s, \omega = \sqrt{1/4 - V_0^2 s^2}$. $\mathcal{F}_t(\nu, \omega), \mathcal{R}$, and \mathcal{S} were given in Ref. [14].

4. Scattering of wave-packets: tunneling, superluminal propagation and causality

In my work with M.S. Marinov we have shown that a description of wave packet propagation simplifies considerably when considered in phase space. The usual analysis in coordinate space is given here first to set the stage for the discussion.

Consider the time evolution of a plane wave having momentum p scattered with the scattering amplitude A(p) and the dispersion $\omega(p)$:

$$|p\rangle \longrightarrow \hat{U}_t^T |p\rangle = \exp[-i\omega(p)t] A(p) |p\rangle$$
 (23)

A wave packet is created from a superposition of plane waves:

$$\Psi^{T}(q;t) = \int dp \,\Phi(p)A(p) \exp[i(pq - \omega t)], \qquad (24)$$

$$\Psi^{F}(q;t) = \int dp \,\Phi(p) \exp[i(pq - \omega t)], \qquad (25)$$

where $\Psi^T(q;t)$ is the scattered wave packet and $\Psi^F(q;t)$ is the freely propagating wave-packet had there been no scattering.

Defining the phase shift $\phi(p)$ of the scattering amplitude as follows:

$$A(p) = |A(p)| \exp[i\phi(p)] , \qquad (26)$$

and applying a stationary phase analysis, one finds the peaks of the freely propagating and scattered wave packets, respectively:

$$q_F^0 = v_g t , \qquad (27)$$

$$q_T^0 = v_g(t - d\phi/d\omega) , \qquad (28)$$

where $v_g \equiv d\omega/dk$ is the well known group velocity and $d\phi/d\omega$ is Wigner phasetime delay. [16] Note, however, that the naive application of the stationary phase argument is correct *only* if A(k) and $\Phi(k)$ are slowly varying. Certainly, A(k) is not slowly varying for deep tunneling, where the phase-time delay is often negative.

It was found that the group velocity can sometimes exceed c and that the phasetime delay can be negative, which gives "superluminal" phenomena or "faster-thanlight" effective velocities. Many works discuss these effects. [17]-[34] The different superluminal phenomena involve no violation of causality.

In the time-independent formulation causality manifests itself in analytical properties of the scattering amplitude: [13]

- $A^*(p) = A(-p^*),$
- $A(p) \to \text{constant as } |p| \to \infty$,
- A(p) is analytic in the upper half of the complex p plane.

How do these properties manifest themselves in real space or in phase space? An argument of causality in coordinate space may assume the following form: define a propagator for scattering in coordinate space in the following way:

$$\Psi^{T}(q;t) = \int dq_0 \langle q | \hat{U}_t^T | q_0 \rangle \Psi(q_0,0), \qquad (29)$$

$$\langle q|\hat{U}_t^T|q_0\rangle = \int dp A(p) \exp\left[ip\left(q-q_0-t\frac{\omega}{p}\right)\right].$$
 (30)

For photons in vacuum (*i.e.*, with no dispersion) $\omega = pc$. Considering the integration over p as a contour integration in the complex p plane and closing the contour in the upper half of the complex plane, one can see that the analytic properties of the amplitude result in the following causal restriction:

$$\omega(p) = cp \longrightarrow \langle q | \hat{U}_t^T | q_0 \rangle = 0 \quad \text{if} \quad q > q_0 + tc .$$
(31)

No information can be transferred faster than the speed of light in vacuum c. Unfortunately, this argument fails when $\omega(p) \neq pc$.

Consider now a similar argument in phase space. The propagator for scattering in the Wigner representation in the elastic channel defined by the scattering amplitude A(p) is given by:

$$\mathcal{T} = \delta(p - p_0) \frac{1}{2\pi} \int d\sigma e^{i\sigma(q - q_0)} e^{-it[\omega(p + \sigma/2) - \omega(p - \sigma/2)]}$$
(32)

$$\times A(\frac{\sigma}{2} + p) A(\frac{\sigma}{2} - p) .$$

An analysis based on an analytic continuation into the complex σ plane reproduces the result for the dispersion relations of photons in vacuum,

$$\omega(p) = cp \longrightarrow \mathcal{T} = 0 \quad \text{if} \quad q > q_0 + tc , \qquad (33)$$

and also gives a new result for massive nonrelativistic particles:

$$\omega(p) = p^2/2m \longrightarrow \mathcal{T} = 0 \quad \text{if} \quad q > q_0 + tp/m \,. \tag{34}$$

No similar condition exists for other dispersion relations, including in particular the relativistic, Klein-Gordon dispersion.

Note that the propagator for free photons, with $\hbar\omega(p) = cp$, is

$$\mathcal{L}_t = \mathcal{T} = \delta(p - p_0)\delta(q_0 + ct - q) , \qquad (35)$$

while the propagator for free massive particles, with $\hbar\omega(p) = p^2/2m$, is

$$\mathcal{L}_t = \mathcal{T} = \delta(p - p_0)\delta(q_0 + tp/m - q) .$$
(36)

5. Scattering from a potential barrier

For one dimensional scattering from a potential barrier, $H = p^2/2m + V(q)$, there are two channels and two amplitudes:

- Transmission amplitude A(p),
- Reflection amplitude B(p).

The evolution kernel for the Wigner function has three parts:

$$\mathcal{L}_t(q, p; q_0, p_0) = \mathcal{T} + \mathcal{R} + \mathcal{S}, \qquad (37)$$

$$\mathcal{T} = \delta(p - p_0) \frac{1}{2\pi} \int d\xi A(p + \frac{\xi}{2}) A(-p + \frac{\xi}{2})$$
(38)

$$\times \exp[i\xi(q-q_0-t\frac{p}{m})],$$

$$\mathcal{R} = \delta(p+p_0) \frac{1}{2\pi} \int d\xi B(p+\frac{\xi}{2}) B(-p+\frac{\xi}{2})$$

$$\times \exp[-i\xi(q+q_0+t\frac{p}{m})],$$
(39)

and contributions from S are exponentially suppressed with time. A general form of the propagator for potential barriers is obtained by closing the contour of the integral and obtaining a sum over the S-matrix singularities, which are simple poles at κ_n , with Im $\kappa_n < 0$.

$$\mathcal{T} = \delta(p - p_0)\delta\left(q - q_0 - t\frac{p}{m}\right) - \delta(p - p_0)\theta\left(q - q_0 - t\frac{p}{m}\right) \\ \times \Sigma_n \operatorname{Re}\left\{C_n(p)\exp\left[i2\left(q - q_0 - t\frac{p}{m}\right)(p - \kappa_n)\right]\right\}.$$
(40)

A simple interpretation of causality in tunneling is obtained: *the barrier removes delayed parts from the freely propagating wave packet*.

The momentum probability distribution of the transmitted part is trivial:

$$\mathcal{P}_t^T(p) = \int dq \rho_t^{W}(\mathbf{q}, \mathbf{p}) = |A(p)|^2 \mathcal{P}_0(p) \,. \tag{41}$$

Using the notation:

$$P_0 = \langle \hat{p} \rangle_0 \; ; \; (\Delta p_0)^2 = \langle (\hat{p} - P_0)^2 \rangle_0 \; , \tag{42}$$

$$Q_0 = \langle \hat{q} \rangle_0 \; ; \; (\Delta q_0)^2 = \langle (\hat{q} - Q_0)^2 \rangle_0 \; , \tag{43}$$

a new result is obtained for the coordinate probability distribution of the transmitted part. An expansion in $(\Delta p_0/P_0)$ gives:

$$\mathcal{P}_{t}^{T}(q) = \frac{1}{2\pi} \int dp \,\rho_{t}^{W}(\mathbf{q}, \mathbf{p}) = \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} n_{l,j} \, N_{l}^{(j)}(q, t) \,, \tag{44}$$

$$N_l^{(j)}(q,t) \equiv \frac{1}{2\pi} \left(\frac{\partial}{\partial q}\right)^j \int_{-\infty}^{\infty} dp (p-P_0)^l \rho_0(q-tp/m,p) \,, \tag{45}$$

$$n_{l,j} \equiv \left(\frac{i}{2}\right)^{j} \sum_{r=0}^{j} \sum_{s=0}^{l} \frac{(-1)^{r}}{r!(j-r)!s!(l-s)!} \times \left[\left(\frac{\partial}{\partial P_{0}}\right)^{(r+s)} A(P_{0}) \right] \left[\left(\frac{\partial}{\partial P_{0}}\right)^{[(j+l)-(r+s)]} A^{*}(P_{0}) \right].$$
(46)

To first order in $(\Delta p_0/P_0)$ this reduces to:

$$\mathcal{P}_t^T(q) \approx |A(P_0)|^2 \left\{ \mathcal{P}_t^f(q) + \frac{P_0}{m} \tau^{\mathrm{W}} \frac{\partial}{\partial q} \mathcal{P}_t^f(q) + \frac{P_0}{m} \tau^{\mathrm{A}} 2M_t^f(q) \right\}, \quad (47)$$

where the 1st derivative and 1st moment of the freely propagating distribution:

$$\mathcal{P}_t^f(q) \equiv \frac{1}{2\pi} \int dp \,\rho_0(q - tp/m, p) , \qquad (48)$$

$$M_t^f(q) \equiv \frac{1}{2\pi} \int dp \, (p - P_0) \rho_0(q - tp/m, p)$$
(49)

are coupled to two real time parameters:

$$\tau^{W} = \frac{m}{P_0} \left[\frac{\partial (\arg A(P_0))}{\partial P_0} \right] = \left[\frac{\partial (\arg A(P_0))}{\partial E} \right], \tag{50}$$

$$\tau^{A} = \frac{m}{P_{0}} \left[\frac{\partial (\ln |A(P_{0})|)}{\partial P_{0}} \right] = \left[\frac{\partial (\ln |A(P_{0})|)}{\partial E} \right] , \qquad (51)$$

which are the real and imaginary parts of the well known complex tunneling time.

As an example, consider an initial Gaussian state:

$$\rho_0(q,p) = C \exp\left[-\frac{1}{2}\left(\frac{q-Q_0}{\Delta q_0}\right)^2\right] \exp\left[-\frac{1}{2}\left(\frac{p-P_0}{\Delta p_0}\right)^2\right] \quad , \tag{52}$$

which for free propagation gives:

$$\mathcal{P}_t^f(q) = \exp\left[-\frac{1}{2}\left(\frac{q-Q}{\Delta q}\right)^2\right] ,$$
 (53)

$$Q \equiv Q_0 + tP_0/m , \qquad (54)$$

$$\Delta q \equiv \sqrt{(\Delta q_0)^2 + (t\Delta p_0/m)^2} , \qquad (55)$$

$$M_t^f(q) = \left(\frac{t(\Delta p_0)^2}{m}\right) \frac{1}{\Delta q} \left(\frac{q-Q}{\Delta q}\right) \mathcal{P}_t^f(q) .$$
 (56)

The transmitted part after scattering is then given by:

$$\mathcal{P}_t^T(q) = |A(P_0)|^2 \mathcal{P}_t^f(q) \left[1 + \tau_0 \frac{P_0}{m} \frac{1}{\Delta q} \left(\frac{q-Q}{\Delta q} \right) \right] , \qquad (57)$$

$$\tau_0 = 2\tau^{\mathcal{A}} \left(\frac{t(\Delta p_0)^2}{m} \right) - \tau^{\mathcal{W}} .$$
(58)

The peak is narrowed and advanced by ΔQ where

$$\Delta Q = \frac{\Delta q \, 2 \, \tau_0 \, P_0/m}{\Delta q + \sqrt{(2\tau_0 \, P_0/m)^2 + (\Delta q)^2}} \quad , \tag{59}$$

$$\Delta Q \approx \tau_0 P_0/m \text{ for } \tau_0 P_0/m \ll \Delta q \quad ,$$
 (60)

$$\Delta Q \approx \Delta q \quad \text{for } \tau_0 P_0/m \gg \Delta q \quad .$$
 (61)

The peak of the scattered wave packet is never advanced more than the width of the freely propagating wave packet. Note that the example of a Gaussian is just an example. The method is general enough to apply to any initial state.

6. Nonclassical energy transfer processes within a single molecule

A transition probability is given by the phase-space overlap integral between initial and final states:

$$\Sigma_{if} = \frac{1}{2\pi} \int dq dp \,\rho_f(q,p) \,\rho_i(q,p) \,. \tag{62}$$

In relaxation processes, a given initial excited state $\hat{\rho}_i$ relaxes into a manifold of all final states with a given energy $\hat{\rho}_f$:

$$\hat{\rho}_f = \Sigma_n |\psi_n(E)\rangle \langle \psi_n(E)| = \delta(E - \hat{H}_f) .$$
(63)

where E is the final energy, and \hat{H}_f is the final state quantum Hamiltonian operator. The transition probability for relaxation processes is thus:

$$\Sigma_{if} = \int dq dp \ \Delta(q, p) \ \rho_i(q, p) = \operatorname{Tr}[\delta(E - \hat{H}_f)\hat{\rho}_i] , \qquad (64)$$

$$\Delta(q,p) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} d\eta \langle q + \frac{\eta}{2} |\delta(E - \hat{H}_f)| q - \frac{\eta}{2} \rangle e^{-ip\eta} \,. \tag{65}$$

In our work we study the integrand:

$$\zeta_{if}(q,p) = \Delta(q,p)\rho_i(q,p), \qquad (66)$$

which in the leading-order semiclassical approximation is:

$$\zeta_{if}^{\rm cl}(q,p) \approx \delta[E - H_{\rm cl}(q,p)]\rho_i(q,p) \,. \tag{67}$$

While the integral gives an estimate for the transition rate, the integrand provides an indication for the preferred channels for the energy. By looking for accepting zones in phase space, *i.e.*, regions of phase space where $\zeta_{if}(q, p)$ is large, we calculate propensity rules and chose between competing channels. The examples we have studied include model potentials of harmonic and nonharmonic oscillators and application to internal conversion in the benzene molecule.

7. Conclusions

Quantum mechanics can be studied in many different representations. The physical results of an experiment or the theoretical predictions for an observable effect do not depend on the representation chosen, but a clever choice often simplifies the analysis and sometimes helps our physical intuition. In this work several cases have been considered where fundamental or complicated problems considerably simplify in the Wigner representation. Applications to atom optics have not been discussed here for lack of space, but are a particularly important experimental field for which these methods may prove useful.

Acknowledgments

I dedicate this paper to the memory of Prof. Michael Marinov. This work was supported by the Israel-USA Binational Science Foundation under grant number 9800460
References

- E.P. Wigner, *Phys. Rev.* 40, 749 (1932). Wigner gives the following footnote: "This expression was found by L. Szilard and the present author some years ago for another purpose".
- [2] J.E. Moyal, Proc. Cambridge Philos. Soc. 45, 99 (1949).
- [3] E.J. Heller, Jour. Chem. Phys. 65 1289 (1976).
- [4] M.V. Berry, Philos. Trans. Roy. Soc. London A 287, 237 (1977).
- [5] H.W. Lee and M.O. Scully, Found. Phys. 13, 61 (1983).
- [6] N.L. Balazs and B.K. Jennings, Phys Rep. 104, 347 (1984).
- [7] M. Hillery, R.F. O'Connell, M.O. Scully, and E.P. Wigner, Phys. Rep. 106, 121 (1984).
- [8] M.S. Marinov, Phys. Lett. A 153, 5 (1991).
- [9] J.G. Muga and R.F. Snider, *Europhys. Lett.* **19**, 569 (1992); R. Sala, S. Brouard, and J.G. Muga, *J. Chem. Phys.* **99**, 2708, (1993); J.G. Muga, R. Sala, and S. Brouard, *Solid State Commun.* **94**, 877 (1995).
- [10] H.W. Lee, Phys Rep. 259, 147 (1995).
- [11] K. Banaszek, K. Wodkiewicz, and W.P. Schleich, *Laser Physics* 10, 123 (2000); A. Czirjaz, R. Kopold, W. Becker, M. Kleber, and W.P. Schleich, *Optics Communications* 179, 29 (2000).
- [12] M. Moshinsky and A. Sharma, Annals of Physics 282, 138 (2000).
- [13] M.S. Marinov and B. Segev, J. Phys. A: Math. Gen. 29, 2839 (1996).
- [14] M.S. Marinov and B. Segev, Phys. Rev. A54, 4752 (1996); 55, 3580 (1997); Found. of Phys. 27, 113-132 (1997).
- [15] B. Segev and E.J. Heller, Journal of Chemical Physics 112, 4004 (2000).
- [16] E.P. Wigner, Phys. Rev. 98, 145 (1955).
- [17] M. Büttiker and R. Landauer, Phys. Rev. Lett. 49, 1739 (1982).
- [18] E. H. Hauge and J. A. Støvneng, Rev. Mod. Phys. 61, 917 (1989).
- [19] C. R. Leavens and G. C. Aers, Phys. Rev. B 39, 1202 (1989).
- [20] A. Ranfagni, D. Mugnai, P. Fabeni, and G. P. Pazzi, *Appl. Phys. Lett.* 58, 774 (1991); A. Ranfagni,
 D. Mugnai, P. Fabeni, G. P. Pazzi, G. Naletto, and C. Sozzi, *Physica B* 175, 283 (1991); A. Ranfagni,
 P. Fabeni, G. P. Pazzi, and D. Mugnai, *Phys. Rev. E* 48, 1453 (1993); L.S. Schulman, A. Ranfagni,
 D. Mugnai, *Physica Scripta* 49, 536 (1994).
- [21] V.S.Olkhovsky and E.Recami, Phys. Rep.214, 339 (1992).
- [22] A. Enders and G. Nimtz, J. de Physique I France 2, 1693 (1992); 1089 (1993); Phys. Rev. B 47, 9605 (1993); Phys. Rev. E 48, 632 (1993).
- [23] J.G. Muga, S. Brouard and R. Sala, *Phys. Lett. A* 167, 24 (1992); *J. Phys.: Condensed Matter* 4, L579 (1992). S. Brouard, R. Sala, and J.G. Muga, *Europhys. Lett.* 22, 159 (1993); *Phys. Rev. A* 49, 4312 (1994).
- [24] R. Y. Chiao, Phys. Rev. A 48(1) R34 (1993).
- [25] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, *Phys. Rev. Lett.* **71**, 708 (1993); A. M. Steinberg and R. Y. Chiao, *Phys. Rev. A* **49**, 3283 (1994); A. M. Steinberg, *Phys. Rev. Lett.* **74**, 2405 (1995); *Phys. Rev. A* **52**, 32 (1995).
- [26] Y. Japha and G. Kurizki, Phys. Rev. A53, 586 (1996).
- [27] Ch. Spielmann, R. Szipocs, A. Sting, and F. Krausz, Phys. Rev. Lett. 73, 2308 (1994).
- [28] Th. Martin and R. Landauer, *Phys. Rev. A* 47, 2023 (1993); R. Landauer and Th. Martin, *Rev. Mod. Phys.* 66, 217 (1994).
- [29] G. Diener, Phys. Lett. A 223, 327 (1996).
- [30] R. Y. Chiao, A. E. Kozhekin and G. Kurizki, Phys. Rev. Lett. 77, 1254 (1996)
- [31] M. W. Mitchell and R. Y. Chiao, Phys. Lett. A 230, 133 (1997).
- [32] A.M.Gabovich and A.I.Voitenko, Phys. Rev. B 55, 1081 (1997).
- [33] Y. Aharonov, B. Reznik, and A. Stern, Phys. Rev. Lett. 81, 2190 (1998).
- [34] 17. B. Segev, P.W. Milonni, J.F. Babb and R.Y. Chiao, Phys. Rev. A 62 (2000).

Geometric Structure of the Big Bang

Michael Heller Vatican Observatory, V-00120 Vatican City State

In the standard approach to the investigation of singularities in general relativity, singularities are treated as points of a "singular boundary" rather than events of spacetime. To treat them as "internal elements" of a given space-time, a generalization of the standard geometric methods is required. A new approach to the singularity problem, based on the noncommutative geometry, is briefly presented. From the results obtained so far an interesting picture of the early universe emerges. In the conceptual framework of noncommutative geometry, a distinction between singular and non-singular states of the universe turns out to be meaningless. "Classical singularities" appear only when the universe passes through the Planck threshold to its commutative phase.

Keywords: general relatitity, Big Bang, cosmology, noncommutative geometry

1. Introduction

The "Big Bang" is rather a popular expression, its geometric counterpart being the "initial singularity." For the mathematician the singularity issue in general relativity constitutes a difficult but challenging problem. There are strong reasons to believe that the mathematical degree of difficulty of this problem reflects the dramatic character of its physical counterpart - the beginning of the Universe. Let us consider the open Friedman universe, whose space extends to infinity. If we contemplate its evolution backwards in time the volume of the universe shrinks, but always remains infinite. To attain zero at the singularity, the volume would have, at one instant, to jump from infinity to zero. This would be both physically and mathematically unacceptable. In view of the above it is clear that — contrary to general opinion — the singularity cannot be regarded as a point in space-time at which the volume of the universe vanishes and the matter density blows up to infinity. The singularity is, rather, a "place" at which the very concept of space-time breaks down. And here we have the problem in all its clarity: how to mathematically determine something which is beyond the model we have at our disposal?

There are essentially two methods to cope with this problem. The first method is to regard singularities as ideal or boundary points of space-time, and to investigate them from within a given space-time by using more or less standard geometric methods. The second approach consists in generalizing the concept of space-time manifold in such a way that singularities could be regarded as "internal elements". Both these methods are, in a sense, complementary. The first, which is a paradigmatic approach in studying classical (*i.e.*, without taking into account quantum gravity effects) singularities in general relativity, is more effective in analyzing concrete singular space-times. The second — still undere development, but already with significant successes – seems to be indispensable in disclosing a source of various "singular situations". Both methods are useful in proving some general theorems concerning singularities.

The goal of the present paper is to describe the second of these methods and its main results. The first approach will be only briefly summarized, in Section 2, to more clearly state the problem and to prepare the stage for further considerations. Section 3 is a brief interlude mentioning an intermediate step which led the present author and his co-workers from the standard approach to the noncommutative modelling of singular space-times. The latter is, in some detail, described in Section 4, and applied to the analysis of the closed Friedman model in Section 5. The main results obtained so far with the help of the noncommutative approach are reviewed in Section 6. And finally, in Section 7, a general image of the "beginning of the universe" is discussed that emerges out of the proposed approach. In the whole of the paper we are more interested in conceptual issues than in technical problems.

2. Space-time model and its breaking down

In general relativity, space-time is modeled by the pair (M, g) where M is a 4 dimensional smooth manifold, and g a smooth Lorentz metric on M with the +2 signature. For the theoretical physicist "smooth" usually means "as smooth as required", and rarely is anything more than C^2 is required. The Lorentz metric on M(one speaks also about the *Lorentz structure*) contains within itself several substructures beautifully "synchronized" with each other, and this artful edifice is exactly what is needed in physics. The total collapse of the space-time structure in the initial singularity means not only breaking down of the space-time stage for physical processes, but also the complete loss of information concerning those aspects of physics which are encoded in space-time geometry (such as: free fall of bodies, speed of light, space and time, gravitational field). A significant breakthrough in coping with the singularity problem was made by Robert Geroch [12, 13] who was able to formulate a clear geometric criterion determining what is meant by breaking down of the space-time structure.

Let $\gamma : I \to M$ be a non-constant geodesic in space-time (M, g). Non-constant geodesic means a geodesic that fails to satisfy the condition: $\gamma(t) = p, p \in M$ for all $t \in I$. The following chain of definitions leads to the Geroch criterion:

- A geodesic γ is complete to the future (to the past) if I = [a,∞] (if I = [-∞, a]), a ∈ R. γ is said to be complete, if I = [-∞,∞] (it can be shown that these definitions are independent of the affine reparametrization of γ).
- Space-time (M, g) is geodesically incomplete (g-incomplete, for short) (to the future, to the past) if in (M, g) there exists at least one incomplete geodesic

(to the future, to the past). If there is no such geodesic, (M, g) is said to be geodesically complete (g-complete).

- Space-time (M, g) is *timelike, null*, or *spacelike g-incomplete (to the future, to the past)* if the g-incomplete geodesic in question is timelike, null or spacelike (to the future, to the past), respectively.
- Correspondingly, one defines timelike, null or spacelike g-completeness of space-time (M, g).

As examples demonstrate, timelike, null and spacelike g - completeness (and g - incompleteness) of space-time are logically independent concepts, *i.e.*, none of these concepts either implies or excludes the others.

Geroch's idea was to regard space-time (M, g) as singularity free if it is timelike and null g-complete, and vice versa, the timelike and null g-incompleteness of spacetime (M, g) is to be regarded as the "minimum condition" for the existence of a singularity, provided that (M, g) is inextendible, *i.e.*, that there is no its smooth isometric embedding into a "larger" space-time (M', g'). This criterion is physically reasonable, since in any timelike or null g-incomplete space -time (M, g) there exists at least one history of a particle or photon which suddenly emerges out of nothing (if (M, g) is incomplete to the past) or disappears into nothingness (if (M, g) is incomplete to the future). In cosmological models with the initial singularity of the Big Bang type all timelike and null geodesics are past incomplete, but spacetimes are also known in which only certain classes of geodesic are (past or future) incomplete.

The above criterion was used by Penrose [31] to prove the first of the series of theorems known as the singularity theorems [17]. Since these theorems did not assume any symmetry postulates, they falsified a so far common belief that singularities in cosmological models were merely by-products of too strong symmetries. The general method in proving singularity theorems consists in combining different kinematic and dynamic conditions so as to obtain the contradiction between these conditions and the assumption of the g-completeness of space-time. In some of the theorems, the assumptions are general enough to be believed to be valid in every universe similar to ours (for more details see [6, 34]).

It was soon realized that the g - incompleteness criterion does not work for all situations which, from the physical point of view, could be regarded as singular. Timelike curves (which are not geodesics) represent histories of nonzero rest-mass particles moving with an acceleration, and if this acceleration is bounded, the motion thus represented is physically realistic, and consequently, space-time should be regarded as singularity-free if it is "complete in the sense of bounded acceleration curves". It was Schmidt [33] who gave this idea an elegant geometric form. He first introduced a *generalized affine parameter* along any curve, and then defined a space-time (M, g) to be *b*-complete (after boundary, see below) if every curve in (M, g)

has infinite length as measured by this parameter. Correspondingly, one speaks of a *b-incomplete* space-time. If a given curve is a geodesic the generalized affine parameter reduces to the usual affine parameter. Every space-time which is b-complete is also g-complete.

The "end-points" of b-incomplete curves were organized by Schmidt into a singular boundary of space-time, called its *b-boundary*. We shall briefly present this construction. Let (M, g) be a space-time, and OM (the connected component) of the orthonormal frame bundle over M, $\pi : OM \to M$, with the Lorentz group SO(3,1) as its structural group. The Levi-Civita connection on M determines the family of Riemann (positive definite) metrics on the total space OM of the frame bundle over M. We select one of these metrics (the further construction does not depend of the particular choice), use it to determine the distance function on OM and, in the usual way, construct the Cauchy completion \overline{OM} of OM. The right ac tion of the group SO(3,1) on OM can be prolonged to \overline{OM} . Now, we define the quotient space $\overline{M} := \overline{OM}/SO(3,1)$ to be the *b-completion* of space-time M. It can be shown that \overline{M} is open and dense in M. We define the *b-boundary of space-time* as $\partial_b M := \overline{M} \setminus M$.

Schmidt's construction was soon commonly accepted as the best available definition of singularities. Unfortunately, however, it was very difficult to effectively compute b-boundaries for concrete space-times. Only a few years later Bosshard [1] and Johnson [26] were able to demonstrate that the b-boundaries of the closed Friedman universe and of the Schwarzschild solution have strongly pathological properties: they are not Hausdorff separated from the rest of space-time and, in both cases, they consist of a single point. This is very dramatic especially as far as the closed Friedman world model is concerned since this model has two singularities – the beginning and the end of the universe. How could they be a single "point"?

There were some attempts to cure the situation (see [10]), but the new proposals were either less elegant than the original construction, or not general enough, and the b-boundary construction began slowly to disappear from scientific literature. One suspects that the source of the above difficulties with singularities is connected with the fact that the methods used to deal with them have been in fact formulated for problems arising within the category of smooth manifolds, whereas space-times with singularities clearly go beyond this category. To cope with stronger types of singularities one must look for more general mathematical methods.

3. Space-times with singularities as structured spaces

Since the work by Koszul [27] it has been known that the geometry of a smooth manifold M can be reconstructed from the algebra $C^{\infty}(M)$ of smooth functions on M. It turns out that it is possible to define a space, more general than a smooth manifold, by repeating Koszul's strategy for any functional algebra (eventually satisfying some additional requirements). Such spaces, usually called *differential spaces*, have

been studied by many authors (for a bibliography of differential spaces see [2]). If one uses a sheaf of functional algebras rather than a single functional algebra, one speaks of *structural spaces*; these have been studied in [22]. We have investigated space-times with various types of singularities in terms of differential and structured spaces (see, [14, 15, 18, 19, 21, 32]), and in particular space-times with malicious singularities [20, 22]. The result is striking!

Let $\overline{M} = M \cup \partial_b M$ be a space-time M with its b-boundary $\partial_b M$. M is open and dense in \overline{M} (\overline{M} is called a b-completed space-time). Let further C be a functional algebra defining M as a differential space. In such a case C(M) is said to be the *differential structure* on M. A prolongation of the differential structure C on M to that of \overline{M} is defined to be an algebra \overline{C} on M such that $\overline{C}(M) = C(M)$. In [20, 22] we have demonstrated that if M is a space-time with at least one malicious singularity in its b-boundary, and $C^{\infty}(M)$ the differential structure on M, then only constant functions can be prolonged to \overline{M} . The same is true if the differential structure on M consists of a sheaf of functional algebras rather than a single algebra. The fact that only constant functions can be prolonged to \overline{M} explains why the spacetime of the closed Friedman world model with its b-boundary collapses to a single point. Indeed, the differential structure of \overline{M} for this model consists only of constant functions, and constant functions do not distinguish points (the value of a constant function at each point is the same). This explains the difficulty, but does not remove it. To go further more powerful methods must be used.

4. Space-time with malicious singularities as a noncommutative space

The differential structures of differential or structured spaces are functional algebras, and as such they are always commutative. It seems natural, in the next step of generalizations, to look for noncommutative (but still associative) algebras. It is the so-called noncommutative geometry that we shall try to use in analysing malicious singularities.

Good introductions to noncommutative geometry are the books by Landi [28] and Madore [29]; one should also consult the monumental work by Connes [7]. Noncommutative spaces often arise when one deals with quotient spaces X/R where X is a space (which can be quite innocuous) and R an equivalence relation. The strategy is to organize X/R into a smooth groupoid (called also a Lie groupoid), and then to consider the C^* -algebra naturally associated with it. If this algebra turns out to be noncommutative one treats it as a noncommutative substitute of the algebra $C^{\infty}(X/R)$. A space defined by this algebra is called a *noncommutative space*.

According to the above strategy, we shall change the b-completion of spacetime M defined as the quotient $\overline{OM}/SO(3,1)$ into a suitable groupoid. The group $\Gamma = SO(3,1)$ acts to the right on the Cauchy completed space \overline{OM} of orthonormal frames over space-time M, $\overline{OM} \times \Gamma \rightarrow \overline{OM}$. This allows us to introduce the

M. Heller

groupoid structure on $G = \overline{OM} \times \Gamma$. Elements of G are pairs of orthonormal frames; they can be represented in the form $\gamma = (p, g)$ (and regarded as an arrow beginning at p and ending at pg). Two elements of G, $\gamma_1 = (p, g_1)$ and $\gamma_2 = (q, g_2)$, can be composed if $q = pg_1$ (if the end of one arrow coincides with the beginning of the second arrow). The inverse of $\gamma = (p, g)$ is $\gamma^{-1} = (pg, g^{-1})$. We define the "set of units" $G^{(0)} = \overline{OM} \times \{e\}$, and two mappings $s, r : G \to G^{(0)}$ by s(p, g) = pg and r(p, g) = p, called the *range* and the *source* mappings, respectively. The set $G^{(2)}$ of composable elements of G is of course

$$G^{(2)} := \{ (\gamma_1, \gamma_2) \in \overline{OM} \times \overline{OM} : s(\gamma_1) = r(\gamma_2) \}.$$

Two elements of G can be composed with each other if they lie in the same fibre $G_p = \pi_{\overline{OM}}^{-1}(p), p \in \overline{OM}$, where $\pi_{\overline{OM}} : G \to \overline{OM}$ is the canonical projection. It can be easily checked that G, structured in this way, satisfies all groupoid axioms.

In what follows two sets are important: the set of all arrows that begin at $p \in \overline{OM}$

$$G^p = \{ (p,g) : g \in \Gamma \},\$$

and the set of all arrows that end at $q \in \overline{OM}$

$$G_q = \{(qg^{-1}, g) : g \in \Gamma\}.$$

Both these sets can be equipped with the structure of the SO(3, 1) manifold. Indeed, these sets can be presented in the form $G^p = \{p\} \times SO(3, 1)$ and $G_q = \{qg^{-1}\} \times SO(3, 1)$, respectively, from which the bijection between these sets and the set SO(3, 1) is evident. With the help of this bijection the manifold structure can be carried out from SO(3, 1) to G^p and G_q . This manifold structure is preserved also if p and q are situated in the singular fiber, i. e., if $p, q \in \overline{\pi}^{-1}(x_0)$ where $x_0 \in \partial_b M$. Of course, the pairs (p, pg) belonging to singular fibres are no longer pairs of orthonormal frames, but rather limits of equivalence classes of pairs of Cauchy sequences of orthonormal frames. From Schmidt's construction it follows that these limits always exist.

Now, one defines the involutive algebra $\mathcal{A} := \mathcal{A}_{const} \oplus \mathcal{A}_c$; where $\mathcal{A}_{const} := \pi_{\overline{OM}}^*(C^{\infty}(\overline{OM}))$ and $\mathcal{A}_c := C_c^{\infty}(G, \mathbb{C})$ is the family of all smooth compactly supported complex valued functions on G. Multiplication "*" in this algebra is defined to be the convolution of functions blonging to \mathcal{A} whenever this definition is meaningful; if it is not, one uses the standard function multiplication. For instance, if $a, b \in \mathcal{A}_c$ then

$$(a * b)(\gamma) = \int_{G_p} a(\gamma_1) b(\gamma_1^{-1} \gamma),$$

for every $\gamma \in G_p$, $p \in \overline{OM}$ (integration is with respect to the left Haar measure); and if $a, b \in \mathcal{A}_{const}$ then

$$(a * b)(\gamma) = a(\gamma) \cdot (\gamma).$$

The involution is defined as

$$a^*(\gamma) = \overline{a(\gamma^{-1})}.$$

The algebra \mathcal{A} , has, for each $p \in \overline{OM}$, a nondegenerate representation $\pi_p : \mathcal{A} \to \operatorname{End}(\mathcal{H})$ in the Hilbert space $\mathcal{H} = L^2(G_p)$, given by

$$(\pi_p(a)\xi)(\gamma) = (a_p * \xi)(\gamma),$$

where a_p is a restricted to the fiber over p. The completion of \mathcal{A} with respect to the norm

$$\parallel a \parallel = \sup_{p \in \overline{OM}} \parallel \pi_p(a) \parallel$$

is a C^* -algebra [7, p. 102] which will be denoted by $C^*(\overline{OM})$. This algebra is regarded as a noncommutative substitute of the functional algebra determining a given space. In this sense, the algebra $C^*(\overline{OM})$ contains all information about space-time M and its b-boundary $\partial_b M$ considered as a noncommutative space.

5. Nonlocal character of singularities

Let M be a smooth manifold. The algebraic counterpart of a point $x \in M$ is the maximal ideal of the algebra $C^{\infty}(M)$ of smooth functions on M consisting of all these functions of $C^{\infty}(M)$ that vanish at x. Noncommutative algebras have, in principle, no such ideals; therefore the concept of point in the noncommutative geometry is, in principle, meaningless. This is also true as far as other local concepts are concerned such as that of a neighborhood of a point.

Let $C^*(\overline{OM})^*$ be the dual of $C^*(\overline{OM})$, *i.e.*, the space of continuous linear functionals on \mathcal{A} with the norm

$$\parallel \omega \parallel = \sup_{a \in C^*(\overline{OM})} \{ |\omega(a)| \colon \parallel a \parallel \le 1 \},$$

for every $\omega \in C^*(\overline{OM})^*$. Each positive ω (*i.e.*, such that $\omega(aa^*) \ge 0$ for all $a \in C^*(\overline{OM})$) with the unit norm is called a *state*. The set of all states is convex; the extremal elements of this set are called *pure states*, the remaining ones – *mixed states*.

Now let \mathcal{A} be the commutative algebra $C_0(V)$ of continuous functions on a compact space V. The states on this algebra are equivalent to a probability measure on V, and one can write

$$\omega_{\mu}(f) = \int f d\mu,$$

for $f \in C_0(V)$. The state ω is a pure state if and only if it is equivalent to the Dirac measure concentrated at a point $x \in V$; in such a case $\omega_x(f) = f(x)$. It is therefore clear that pure states can be identified with points of V, and the algebra \mathcal{A} can be regarded as an algebra of functions defined on them. Also in the case

of a noncommutative space one can regard pure states of the corresponding noncommutative algebra as generalizations of the usual concept of point.

Let us return to a noncommutative C^* -algebra \mathcal{A} , and let π be its representation in a Hilbert space \mathcal{H} , let also $\xi \in \mathcal{H}$. In such a case, $a \mapsto (\pi(a)\xi, \xi)$, $a \in \mathcal{A}$, is a positive form on \mathcal{A} . This form is a pure state if and only if π is a nonzero irreducible representation of \mathcal{A} in \mathcal{H} . Let further π_1 and π_2 be two representations of the algebra \mathcal{A} in two Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 , correspondingly. The representations π_1 and π_2 are said to be *equivalent* representations of \mathcal{A} if there is an isomorphism between \mathcal{H}_1 and \mathcal{H}_2 such that $\pi_1(a) = \pi_2(a)$ for every $a \in \mathcal{A}$.

Let us apply the above conceptual machinery to the space-time of the closed Friedman model with its b-boundary regarded as a noncommutative space. The initial and final singularities are two distinct structures given by two representations (strictly speaking by two equivalence classes of representations, each consisting of only one element)

$$\pi_{p_i}: C^{\infty}{}_c(G, \mathbf{C}) \to \operatorname{End} L^2(G_{p_i}),$$

i = 1, 2, where p_1 is the single "limit frame" in the singular fibre over the initial singularity, and p_2 is the single "limit frame" in the singular fibre over the final singularity. Correspondingly, the two singularities can be given by two states $s \mapsto (\pi_{p_i}\xi,\xi), s \in C^*(\overline{OM}), \xi \in L^2(G_{p_i}) i = 1, 2.$

6. Emergence of singularities

As should be expected, the algebra $\mathcal{A} = C^*(\overline{OM})$ contains the information about space-time and its singularities. In this sense, we shall speak about the space-time M associated with the algebra $\mathcal{A} = C^*(\overline{OM})$. In [23] we have proved several theorems which give a nice overview of the emergence of singularities in various situations. We shall quote these results without proofs, but first let us introduce two useful concepts.

We define the following subalgebra of \mathcal{A}

$$\mathcal{A}_{proj} := pr^* C^\infty(M, \mathbf{C}),$$

where $pr = \pi_{\overline{M}} \circ \pi_{\overline{OM}} \to \overline{M}$, is the obvious projection. The subalgebra \mathcal{A}_{proj} consists of functions which are constant on the equivalence classes of fibres of G where two fibres G_p and G_q , $p, q \in \overline{OM}$, are equivalent, if there exists $g \in \Gamma$ such that q = pg. We evidently have $\mathcal{A}_{proj} \subset \mathcal{A}_{const}$.

We also introduce the family of Γ -invariant functions $\mathcal{A}_{\Gamma} \subset \mathcal{A}$, *i.e.*, the family of functions of \mathcal{A} that are constant on the orbits of the action of Γ .

Let us remember that *regular singularities* are those which originate from cutting off some parts of a space-time, and *quasi-regular singularities* are those which originate essentially from cutting off some parts of a space-time and gluing the resulting edges together (for details of the singularity classification see [11]). Now, we can summarize our main results:

- In the space-time associated with the algebra A there is no singularity if and only if A_{proj} ≃ C[∞](M, C).
- 2. The space-time M associated with the algebra \mathcal{A} contains at least one malicious singularity if and only if $\mathcal{A}_{proj} \simeq \mathbf{C}$.
- 3. In the space-time M associated with the algebra \mathcal{A} there is an elementary quasiregular singularity (but there are no stronger singularities) if and only if there exists a discrete group Γ_0 of isometries of M such that $\mathcal{A}_{proj} \simeq C^{\infty}(M)_{\Gamma_0}$.
- 4. In the space-time M associated with the algebra A there is a regular singularity (but there are no stronger singularities) if and only if the groupoid G = OM × Γ is a subspace of a "larger" groupoid G = E × Γ, where OM is a subspace (of constant dimension in the sense of Sikorski) of the space E. In such a case A_{proj} is a localization of Ā_{proj} to G; Ā_{proj} is here the subalgebra of projectible functions on G, i.e., A_{proj} = (Ā_{proj})_G where (Ā_{proj})_G is the algebra of complex valued functions on G which are local restrictions of functions belonging to (Ā_{proj}).

Proofs of these statements can be found in [23]. In agreement with the nonlocal character of the noncommutative algebra \mathcal{A} , the above theorems convey the information about the structure of space-times with singularities rather than about the structure of singularities themselves. Let us notice that if in a given space-time there are singularities of various kinds, the strongest singularity determines the structure of the algebra \mathcal{A} . Regular singularities are very mild singularities (they can hardly be called singularities), they do not change the family \mathcal{A}_{proj} but only narrow its domain.

7. Big Bang and quantum cosmology

The algebra \mathcal{A} , encoding in itself the information about the structure of space-time with singularities, is nonlocal and, consequently, singularities cannot be regarded as points in space-time. However, we can meaningfully speak of pure states of the algebra \mathcal{A} . Each of them is represented by an operator algebra in a Hilbert space, and there is no distinction between singular and nonsingular states. This means that, in the noncommutative setting, the question on the existence or nonexistence of singularities does not even arise.

Is this mathematical formalism only an artificial tool to deal with classical singularities, or could it also somehow reflect physics of the quantum gravity regime? The fact that the states on the algebra \mathcal{A} are represented as operator algebras in a Hilbert space (a typically quantum structure!) could be a hint that the above presented mathematical formalism is indeed somehow related to quantum phenomena in the early universe. In fact, there are several attempts to create a quantum gravity theory based on noncommutative geometry (see, for instance, [3, 4, 5, 8, 24, 25, 30]). However, the following discussion is independent of any of these. We shall simply explore some consequences of the assumption that the algebra A contains information about the pre-Planck era of the universe.

As we have seen, the algebra \mathcal{A} can be completed to the C^* -algebra. This is important because C^* algebras are standard tools in the quantization of physical fields. Within the noncommutative framework, C^* algebras also generalize the standard concept of topology, and the generalization is so powerful that even non-Hausdorff cases can be dealt with by using this method (see [7, p. 79]). This could provide a mathematical basis for a noncommutative version of a "topological foam" supposedly reigning in the quantum gravity regime. However, this version of the idea is much more radical than, for example, the one developed by Hawking (see, *e.g.*, [16]). It is not even a "foamy space-time" that we meet here, but rather a situation in which there is no space and no time in the usual meanings of these terms. In spite of this fact, there could be a true dynamics in the noncommutative regime; for instance, dynamical equations could be written in terms of derivations of the algebra \mathcal{A} (see [9, 24]).

The transition from the noncommutative regime to the usual space-time geometry can be thought of as a kind of "phase transition"; mathematically it corresponds to the transition from the noncommutative algebra \mathcal{A} to its center $\mathcal{Z}(\mathcal{A})$ (or to $\mathcal{A}_{proj} \subset \mathcal{Z}(\mathcal{A})$). In this way, the usual space-time M together with it its singular boundary $\partial_b M$ (*i.e.*, with its singularities) is recovered. It is supposed that this happens when the universe passes trough the Planck threshold. Of course, the same can be - mutatis mutandis - said about final singularities, for instance in the closed Friedman world model or in the gravitational collapse of a massive object (the Schwarzschild singularity is also malicious), but let us focus on the "Big Bang philosophy". We are confronted here with the completely new situation. So far people believed that there are only two possibilities: either the future quantum gravity theory will remove the initial singularity from the cosmological model, or not. If the proposal discussed in the present work is true, there is the third possibility. On the fundamental level, beyond the Planck threshold there is no distinction between singular and nonsingular states of the universe, and the question concerning the existence or nonexistence of the initial singularity is meaningless. The singularity is produced in the process of the formation of macroscopic physics, when space-time emerges from the quantum "foam" (geometry of this process has been studied in [23]). Consequently, it is only from the perspective of the macroscopic observer that the question about the "beginning of the universe" (and possibly about its "end") becomes meaningful. After all, if space-time is a macroscopic concept, its breaking down — the singularity — should also be a macroscopic catastrophe.

References

- [1] Bosshard B.: Commun. Math. Phys. 46, 263-268, 1976.
- [2] Buchner K., Heller M., Multarzyński P., Sasin W.: Acta Cosmol. (Cracow) 19, 111-129, 1993.
- [3] Chamseddine A. H., Connes A.: Phys. Rev. Lett. 24, 4868-4871, 1996.
- [4] Chauseddine A. H., Felder G., Frölich J.: noncommutative geometry. Commun. Math. Phys. 155, 205-217, 1993.
- [5] Chamseddine A. H., Frölich J., Grandjean O.: J. Math. Phys. 36, 6255-6275, 1995.
- [6] Clarke C. J. S.: The Analysis of Space-Time Singularities, Cambridge University Press, 1993.
- [7] Connes A.: Noncommutative Geometry, Academic Press, New York London, 1994.
- [8] Connes A.: Commun. Math. Phys. 182, 155-176, 1996.
- [9] Connes A., Rovelli C.: Class. Quantum Grav. 11, 2899-2917, 1994.
- [10] Dodson C. T. J.: Int. J. Theor. Phys. 17, 389-504, 1978.
- [11] Ellis G. F. R., Schmidt B. G.: General Rel. Grav. 11, 915-953, 1977.
- [12] Geroch R.: Ann. Phys. (New York) 48, 526-540, 1968.
- [13] Geroch R.: J. Math. Phys. 9, 450-465, 1968.
- [14] Gruszczak J., Heller M.: Int. J. Theor. Phys. 32, 625-648, 1993.
- [15] Gruszczak J., Heller M., Sasin W.: Acta Cosmol. (Cracow) 18, 45-55, 1992.
- [16] Hawking S. W.: The path-integral approach to quantum gravity, in General Relativity An Einstein Centenary Survey, ed. S. W. Hawking, W. Israel, Cambridge University Press, Cambridge, 1979.
- [17] Hawking S. W., Ellis G. R. S.: The Large Scale Structure of Space-Time Cambridge University Press, 1973.
- [18] Heller M., Multarzyński P., Sasin W.: Acta Cosmol. (Cracow) 16, 153-85, 1989.
- [19] Heller M., Sasin W.: Acta Cosmol. (Cracow), 17, 7-18, 1991.
- [20] Heller M., Sasin W.: Gen. Rel. Grav. 26, 797-811, 1994.
- [21] Heller M., Sasin W.: Acta Cosmol. (Cracow) 21, 47-60, 1995.
- [22] Heller M., Sasin W.: J. Math. Phys. 36, 1995, 3644-3662.
- [23] Heller M., Sasin W.: Gen. Rel. Grav. 31, 555-570, 1999.
- [24] Heller M., Sasin W.: Int. J. Theor. Phys. 38, 1619-1642, 1999.
- [25] Heller M., Sasin W., Lambert D.: J. Math. Phys. 38, 5840-5853, 1997.
- [26] Johnson R. A.: J. Math. Phys. 18, 898-902, 1977.
- [27] Koszul J. L.: Fibre Bundles and Differential Geometry, Tata Institute of Fundamental Research, Bombay, 1960.
- [28] Landi G.: An Introduction to Noncommutative Spaces and Their Geometries, Springer, Berlin Heidelberg, 1997.
- [29] Madore J.: An Introduction to Noncommutative Differential Geometry and Its Physical Applications, 2nd. ed., Cambridge University Press, Cambridge, 1999.
- [30] Madore J., Mourad J.: J. Math. Phys. 39, 423-442, 1998.
- [31] Penrose R.: Phys. Rev. Lett. 14, 57-59, 1965.
- [32] Sasin W.: Differential spaces and singularities in differential space-times, Demonstratio Math. (Warsaw) 24, 601-634, 1991.
- [33] Schmidt B. G.: Gen. Rel. Grav. 1, 269-280, 1971.
- [34] Tipler F. J., Clarke C. J. S., Ellis G. F. R.: Singularities and horizons A review article, in General Relativity and Gravitation, vol. 2, ed. Held A., Plenum Press, New York London, 1980, 97-206.

The Quantities c^4/G and c^5/G and the Basic Equations of Quantum Mechanics

L. Kostro

Department for Logic, Methodology and Philosophy of Science, University of Gdańsk,

ul. Bielańska 5, 80-851 Gdańsk, Poland

The quantities (e^4/G) and (e^5/G) when introduced into the classical equations of Newton and Coulomb have the meaning of the maximum force and the maximum power, but when we introduce them into the basic equations of Quantum Mechanics we do not see their physical meaning clearly.

Keywords: special relativity, equations of Quantum Mechanics

1. Introduction

In two recent papers [1-2] the quantity c^4/G was interpreted as the greatest possible force in Nature. In the third paper [3], following I.R. Kenyon [4], the quantity c^5/G was interpreted as the greatest possible power. In the three above papers I have limited myself to classical considerations. I have shown, *e.g.*, that the classical Newton law and the classical Coulomb law can be rewritten in the following way: Newton force

Newton

$$F_N = Gm^2/R^2 = (c^4/G)(l_G^2/R^2)$$
 when $m_1 = m_2$,

and

$$F_N = Gm_1m_2/R^2 = (c^4/G)(l_{G1}l_{G2}/R^2)$$
 when $m_1 \neq m_2$.

Coulomb force

 $F_C = KQq/R^2 = KZ_1 eZ_2 e/R^2 = (c^4/G)(l_s^2/R^2)(Z_1Z_2)$

I also indicated that the quantities c^4/G and c^5/G and their inverses appear in the equations of General Relativity [1,2,3], and Kenyon has given his interpretation [4] of this fact.

In my considerations I use the following constants and constant coefficients: c — velocity of light in vacuum; G the gravitational constant; \hbar Planck's constant; e — the elementary electrical charge; m — the mass of an elementary particle; $K = 1/4\pi\varepsilon_0$. I take into account also the units of length, time and mass determined by the following set of constants (c,G,m), (c,G,e), (c,G,\hbar) , (c,G,g_{Str}) , (c,G,g_W) , where m, e, g_{Str}, g_W are the charges of four fundamental interactions, respectively. Using dimensional analysis we obtain the following units:

1. gravitational length l_G , time t_G , and mass m_G :

$$l_G = Gm/c^2; \quad t_G = Gm/c^3; \quad m_G = m$$

2. J.G. Stoney's length l_S , time t_S , and mass m_S introduced by him in 1874 [5-6]

$$l_S = (KGe^2/c^4)^{1/2};$$
 $t_S = (KGe^2/c^6)^{1/2};$ $m_S = (Ke^2/G)^{1/2}.$

3. M. Planck's length l_P , time t_P , and mass m_P introduced by him in 1899 [7]

$$l_P = (\hbar G/c^3)^{1/2}; \quad t_P = (\hbar G/c^5)^{1/2}; \quad m_S = (\hbar c/G)^{1/2}.$$

4. Length l_{Str} , time t_{Str} and mass m_{Str} connected with the strong interactions

$$\begin{split} l_{Str} &= (1/4\pi \ G \ g_{Str}^2/c^4)^{1/2}; \quad t_{Str} = (1/4\pi \ G \ g_{Str}^2/c^6)^{1/2}; \\ m_{Str} &= (1/4\pi \ g_{Str}^2/G)^{1/2}. \end{split}$$

5. Length l_W , time t_W , and mass m_W connected with the weak interactions

$$\begin{split} l_W = &(I/4\pi\;G\;g_W^2/c^4)^{1/2}; \quad t_W = (I/4\pi\;G\;g_W^2/c^6)^{1/2}; \\ m_W = &(I/4\pi\;g_W^2/G)^{1/2}. \end{split}$$

It is interesting to note that forces F and powers P connected with these units are all equal:

$$F_G = F_S = F_P = F_{Str} = F_W = c^4/G = 1.2107x10^{44}N,$$

$$P_G = P_S = P_P = P_{Str} = P_W = c^5/G = 3.63x10^{52}W.$$

2. The quantities c⁴/G and c⁵/G and Einstein's Principle of mass and energy equivalence

It is interesting to note that Einstein's Principle of mass and energy equivalence $E = mc^2$ can be rewritten in the following way:

$$E = mc^2 = (c^4/G)l_G = (c^5/G)t_G.$$

This fact shows, perhaps once again, the dynamical nature of the matter. If an elementary particle could deliver its total energy $E = mc^2$ acting on the path equal to l_G during time t_G , then it could denote the particle's greatest force (c^4/G) and power (c^5/G) . If this could happen then the maximum force (c^4/G) and maximum power (c^5/G) would be hidden in every particle. Perhaps in the future mankind will find the circumstances in which this is possible. At the present time, however, we can only dare to interpret the two quanties as maxima.

3. The quantities $c^{4/G}$ and $c^{5/G}$ and Schrödinger equation

As is well known, the Schrödinger equation is the basic equation of non relativistic Quantum Mechanics. In textbooks, it is written in the following way

 $(-\hbar^2/2m)[\partial^2\psi(x,t)/\partial x^2] + V(x,t)\,\psi(x,t) = i\hbar[\partial\psi(x,t)/\partial t],$

where $V = -Ke^2/r$ is the Coulomb potential. In these equations we find the constants \hbar, m, e and the coefficient K. Since the constants used in physics and the units determined by them are correlated and interconnected, it is therefore not difficult to rewrite the Schrödinger equation in such a way that the quantities c^4/G and c^5/G and the considered units appear in it.

The Coulomb potential can be rewritten as follows

 $V=-Ke^2/r=-(c^4/G)(l_s^2/r)=-(c^5/G)(l_St_S/r)$ and the Schrödinger equation in the following way

$$-(\hbar/2m)(c^4/G)(l_P t_P)[\partial^2 \psi(x,t)/\partial x^2] - [(c^5/G)(l_S t_S/r)](x,t)\psi(x,t) = i(c^4/G)(l_P t_P)[\partial \psi(x,t)/\partial t].$$
(1)

Since $l_S t_S = l_P t_P \alpha$ (where $\alpha = Ke^2/\hbar c$ is the fine structure constant) we obtain also

$$-(\hbar/2m)(c^4/G)(l_P t_P)[\partial^2 \psi(x,t)/\partial x^2] -$$

$$[(c^5/G)(l_P t_P \alpha/r)](x,t)\psi(x,t) = i(c^4/G)(l_P t_P)[\partial \psi(x,t)/\partial t].$$
(2)

As we can see, in the Schrödinger equation written in this way, there appear not only the quantities c^4/G and c^5/G but also the Planck length and time and Stoney's length and time. We see also that Planck's constant \hbar is related to the quantities c^4/G and c^5/G as follows:

$$\hbar = (c^4/G)(l_P t_P) = (c^5/G)t_P^2$$

When we divide both sides of eq. (2) by $(l_P t_P)$ we obtain

$$-(c^{4}/G)(\hbar/2m)[\partial^{2}\psi(x,t)/\partial x^{2}] - [(c^{5}/G)(\alpha/r)](x,t)\psi(x,t) = i(c^{4}/G)[\partial\psi(x,t)/\partial t].$$
(3)

We must be aware, however, that this division changes the numerical value and dimensions of both sides of the equation. The three dimensional Hamilton operator

 $H = -(\hbar^2/2m) \left[(\partial^2 \psi/\partial x^2) + (\partial^2 \psi/\partial y^2) + (\partial^2 \psi/\partial z^2) \right] + V$ can be rewritten by introducing c^4/G and c^5/G as follows

$$H =$$

$$-(\hbar/2m)(c^4/G)(l_P t_P) \left[(\partial^2 \psi/\partial x^2) + (\partial^2 \psi/\partial y^2) + (\partial^2 \psi/\partial z^2)\right] + V$$

When

$$V = -Ke^2/r = -(c^4/G)(l_s^2/r) = -(c^5/G)(l_S t_S/r),$$

then we can write

H =

$$-(\hbar/2m)(c^4/G)(l_P t_P)[(\partial^2 \psi/\partial x^2) + (\partial^2 \psi/\partial y^2) + (\partial^2 \psi/\partial z^2)] - (c^5/G)(l_S t_S/r).$$

4. The quantities c^{4}/G and c^{5}/G and the Klein-Gordon equation

We now consider the Klein-Gordon equation, written, e.g., for the π -mesons.

$$-\hbar^{2}(\partial^{2}\psi/\partial t^{2}) = -\hbar^{2}c^{2}(\partial^{2}\psi/\partial x^{2} + \partial^{2}\psi/\partial y^{2} + \partial^{2}\psi/\partial z^{2}) + m_{\pi}^{2}c^{4}\psi.$$

$$\tag{4}$$

Taking into consideration the quantities c^4/G and c^5/G , the Klein-Gordon equation can be rewritten:

$$-\hbar (c^4/G)(l_P t_P)(\partial^2 \psi/\partial t^2) = -\hbar c (c^5/G)(l_P t_P)(\partial^2 \psi/\partial x^2 + \partial^2 \psi/\partial y^2 + \partial^2 \psi/\partial z^2) + m_{\pi} c^2 (c^4/G) l_G \psi.$$
(5)

As we can see the Planck charge raised to the second power $\hbar c$ is related to (c^5/G) as follows

 $\hbar c = (c^5/G)(l_P t_P).$ Since $m_\pi c^2 = (c^4/G)l_{G\pi} = (c^5/G)t_{G\pi}$ the eq. (5) can be also written as follows

$$-\hbar (c^4/G)(l_P t_P)(\partial^2 \psi/\partial t^2) = -\hbar c (c^5/G)(l_P t_P)(\partial^2 \psi/\partial x^2 + \partial^2 \psi/\partial y^2 + \partial^2 \psi/\partial z^2) + (c^4/G)(c^5/G)l_{G\pi}t_{G\pi}\psi.$$
(6)

Since $l_{G\pi}t_{G\pi} = (l_P t_P)\alpha_{G\pi}$ (where $\alpha_{G\pi} = Gm_{\pi}^2/\hbar c$ is the coupling constant of gravitational interactions between two particles of the same mass, in our case the coupling constant of gravitational interactions between two mesons π). Eq. (6) can be rewritten as follows

$$-\hbar (c^4/G)(l_P t_P)(\partial^2 \psi/\partial t^2) =$$

$$-\hbar c (c^5/G)(l_P t_P)(\partial^2 \psi/\partial x^2 + \partial^2 \psi/\partial y^2 + \partial^2 \psi/\partial z^2) +$$

$$(c^4/G)(c^5/G)(l_P t_P)\alpha_{G\pi}\psi.$$
(7)

Dividing both sides of eq. (7) by $(l_P t_P)$ we obtain

$$-\hbar(c^4/G)(\partial^2\psi/\partial t^2) = -\hbar c(c^5/G)(\partial^2\psi/\partial x^2 + \partial^2\psi/\partial y^2 + \partial^2\psi/\partial z^2) + (c^4/G)(c^5G)\alpha_{G\pi}\psi$$

We must be aware, however, that this division changes the numerical value and dimensions of both sides of the equation.

5. The quantities c^{4}/G and c^{5}/G and the Dirac equation

The Dirac equation can be written as follows:

$$\gamma^{\mu}(\hbar/i)(\partial\psi/\partial x^{\mu}) + mc\psi = 0,$$

where the matrices γ_{μ} have the following properties

$$egin{aligned} &\gamma_\mu\gamma_
u+\gamma_
u\gamma_\mu=g_{\mu
u}, \ &\gamma_o=\gamma_o^\dagger, \ \gamma_i=-\gamma_i^\dagger. \end{aligned}$$

where † means the hermitean conjugation.

Taking the quantities c^4/G and c^5/G into consideration, the Klein-Gordon equation can be rewritten:

 $\gamma^{\mu}[(c^5/G)t_p^2/i)](\partial\psi/\partial x^{\mu}) + (c^4/G)t_G\psi = 0,$ where $t_P = (\hbar G/c^5)^{1/2}$ is the Planck time and $t_G = Gm/c^3$ is the gravitational time.

6. Conclusion

It was very easy to introduce the quantities (c^4/G) and (c^5/G) into the basic equations of Quantum Mechanics (and we might even say that introducing them constitutes a very trivial operation), but it is very difficult to interpret the role they play in these equations. When we introduce the quantities (c^4/G) and (c^5/G) , e.g., into the classical equations of Newton and Coulomb, we immediately see their meaning as limiting quantities, *i.e.*, the maximum force and the maximum power, but when we introduce them into the basic equations of Quantum Mechanics we do not see their physical meaning clearly. They do, however, work in these equations, and therefore we might imagine that they play a role as constants as limiting quantities here also: the maximum force and the maximum force.

References

- [1] L. Kostro, B. Lange, Physics Essays Vol. 10, No 1, March 1999.
- [2] L. Kostro, B. Lange, The force c⁴/G and Relativity Theory, in *Physical Interpretations of Relativity Theory*, (Proceedings ed. M.C. Duffy), British Society for Philosophy of Science, Imperial College, London 11-14 Sept.1998 pp.183-193.
- [3] L. Kostro, B. Lange, The power c⁵/G and Relativity Theory, in *Physical Interpretations of Relativity Theory*, (Later Papers ed. M.C. Duffy), British Society for Philosophy of Science, Imperial College, London 1998.
- [4] I.R. Kenyon, General Relativity, (Oxford University Press 1990).
- [5] G.J. Stoney, Phil. Mag. 5, 381 (1881).
- [6] J.D. Barrow, Q.J. R. Astr. Soc., 24, 24 (1983).
- [7] M. Planck, Secundogeniture d. Preus. Akad. Wiss. /Mitteilung/5, 440 (1899), and in Ann. d. Phys. 1, 69 (1900).

Stability of the Bell-shaped Excitations in Discrete Models of Molecular Chains

Paweł Machnikowski, Piotr Magnuszewski, and Andrzej Radosz Wrocław Technical University ul.Wybrzeze Wyspiańskiego 27, 50-370 Wrocław, Poland

It is shown that the bell-shaped solitary waves in the asymmetric ϕ^4 field model are unstable and correspond to the saddle points of the potential energy. In the discrete model, the potential energy becomes rough: bell-shape configurations may appear stable.

Keywords: molecular chains, soliton solutions, stability

1. The model

In this paper we study the stability properties of a model for a one-dimensional molecular chain. The model is described by the Hamiltonian

$$H = \sum_{n} \left[\frac{1}{2} \dot{u}_{n}^{2}(t) + \frac{k}{2} (u_{n+1}(t) - u_{n}(t))^{2} + U(u_{n}(t)) \right],$$
(1a)

which in the continuum limit transforms into

$$H = \int \mathrm{d}x \left[\frac{1}{2} \dot{u}^2(x,t) + \frac{1}{2} u'^2(x,t) + U\left(u(x,t)\right) \right]. \tag{1b}$$

This model is believed to reflect certain important properties of real systems: both strongly anisotropic 3-dimensional ones and strictly one-dimensional ones. An example of the latter is the DNA double helix (Fig. 1).

A system with a symmetric double-well potential U(u) (Fig. 2a) appears in a variety of applications, and its dynamics has been extensively investigated [1, 2, 3, 4, 5]. Two forms of the on-site potential have been widely used: the ϕ^4 potential

$$U_{\phi^4}(u) = \frac{1}{4}u^4 - \frac{1}{\sqrt{2}}u^3 + \frac{1}{2}u^2,$$
(2)

and the double-Morse potential

$$U_{\rm DM}(u) = \frac{1}{2} U_0 \left\{ A \cosh\left[\alpha(u - u_0)\right] - 1 \right\}^2.$$
(3)

It is known that this model supports localized topological excitations, kinks (Fig. 2b), described in the case (2) as

$$u(x,t) = \frac{1}{\sqrt{2}} \left[1 + \tanh(\gamma \frac{x - vt}{2}) \right],$$

Extensions of Quantum Physics

edited by A. Horzela and E. Kapuścik (Montreal: Apeiron 2002)

121



Fig. 1. The DNA double helix - an example of a molecular chain.

which not only are stable but also to a certain extent preserve their identity in collisions. The effect of discreteness has also been studied for this model and for the closely related sine-Gordon model [2, 6, 7, 8, 9, 10, 11, 12] leading to the accurate description of the effective Peierls–Nabarro potential acting on a discrete kink and of the related kink trapping and radiation effects.

Less work has been devoted to the asymmetric case

$$U(u) = \frac{1}{4}u^4 - \frac{B}{3}u^3 + \frac{1}{2}u^2,$$

(Fig. 3a; $B_0 = 3/\sqrt{2}$ corresponds to the symmetric model) which is also important in various applications [13, 14, 15, 16, 17]. Gordon showed that in this system, stationary, localized bell-shaped solitary waves (Fig. 3b) exist [18, 19],

$$u(x,t) = rac{a}{b + \cosh[\gamma(x - vt)]}.$$



Fig. 2. The symmetric potential and the corresponding kink solution.



Fig. 3. The asymmetric on-site potential and its solutions for various degrees of asymmetry.

The same author has derived topological (kink) solutions which may move with constant velocity if damping is present [20, 21]. Xu and Huang [22] and Xu and Zhou [23] have shown that in the continuum limit in a two-component model, where the anharmonic system (2) is coupled in a special way to another, harmonic system, the equations of motion may be reduced to a single-field problem and solved (see also [24]). The bell-shaped solution has been proposed as a transport mechanism in molecular chains.

It has also been suggested that metastable configurational states are important for the conformational dynamics of the DNA macromolecule [25, 26].

Whereas recently we have shown [27] that the bell-shaped solution is unstable in the asymmetric system in the continuum limit, thorough investigation suggested that the discreteness might be essential in this case. In fact, it appears that this sort of excitation, corresponding in the continuum limit to a saddle point of the potential energy, becomes stabilized in the discrete lattices.

2. Stability of the bell-shaped solitary wave

It is clear that the energy of the bell-shaped solution is higher than the energy of both the false vacuum state, u = 0, and the true vacuum state, $u = u_{\min}$. Since the bell-shaped solution is a non-topological solution, it may be continuously deformed to any of these two states. Therefore, it should not be expected to be stable [28, 29]. Indeed, let us follow the standard linear stability analysis [3, 28] and write in the bell-shaped solution's resting frame, $x = \gamma(\xi - v\tau)$, $s = \gamma(\xi + v\tau)$,

$$\varphi(x,s) = \varphi_{\rm b}(x) + \phi(x) {\rm e}^{i\omega s}.$$
(4)

The equation of motion linearized in ϕ takes the form of a Schrödinger–like eigenvalue problem

$$-\partial_x^2 \phi(x) + V(x)\phi(x) = \omega^2 \phi(x), \tag{5}$$

where

$$V(x) = \left[\frac{\mathrm{d}^2 U(\varphi)}{\mathrm{d}\varphi^2}\right]_{\varphi = \varphi_{\mathrm{b}}(x)} = 3\varphi_{\mathrm{b}}^2(x) - 2B\varphi_{\mathrm{b}}(x) + 1.$$
(6)

The characteristic shape of this potential is shown in the Fig. 4. The plot corresponds to a nearly degenerate potential, where the bell-shaped solitary wave separates into two kinks. Each of these generates a Pöshl–Teller well in the stability potential (6).

The function

$$\phi_{\rm G}(x) = \partial_x \varphi_{\rm b}(x) \tag{7}$$

is a solution of (5) corresponding to $\omega^2 = 0$. This is a characteristic excitation related to broken translational invariance: all of the bell-shape positions along the X axis correspond to the same value of energy. Had the broken continuous symmetry been an internal one this excitation would have been a gapless Goldstone boson; because the broken symmetry is translational one but not an internal one, a pseudo-Goldstone, separated, zero-frequency mode is observed. In the case of the symmetric system, the pseudo-Goldstone mode, $\omega^2 = 0$, is nodeless, corresponding to the lowest-energy perturbation of the kink. However, for a bell-shaped solution φ_b , the function (7) has one node. Therefore, there is a ground state solution to (5) belonging to a lower eigenvalue, $\omega^2 < 0$. Such a solution, as is clear from (4), has imaginary frequency, *i.e.*, it explodes exponentially, destroying the original solution. For such a solution, $|\omega|^{-1}$ may be interpreted as the lifetime of the bell-shaped solitary wave. In the almost degenerate case (the potential in the Fig. 4) the negative eigenvalue may be obtained by semiclassical methods.



Fig. 4. The potential for stability analysis for almost degenerate minima of the on-site potential, $B = B_0 + 10^{-6}$.

3. Stability in discrete models

In the discrete model there is no translational continuous symmetry: at various positions along the chain the kink or bell-shaped configurations have different potential energy. In the case of a kink, it is possible to determine this potential energy in a unique way and to define the so-called Peierls–Nabarro potential.

Replacing the continuum translational symmetry with the discrete one results in shifting the zero-frequency Goldstone mode up. As a consequence, it turns into a positive-frequency oscillatory mode associated, in the symmetric case, with kink oscillation around the minimum of the Peierls–Nabarro potential. In the asymmetric case the corresponding solution – the bell-shape – has both the zero-frequency mode and one imaginary-frequency mode. One might expect that for a strong enough discreteness this exploding mode will also be shifted up enough to become an oscillatory one. Actually, numerical analysis of the small vibration spectrum around a bell-shaped configuration confirms this expectation. Stabilization may also be expected on the grounds of the following argument.

In the limit of independent oscillators $(k \rightarrow 0)$, any configuration with some nodes in the right and some in the left well is stable. In particular any bell-shaped configuration, *i.e.*, one with a certain number of consecutive nodes placed exactly in the global minimum is stable. One may expect that for low values of k, all these configurations survive in a slightly changed form. Fig. 5 shows one of the stable bell-shaped configurations. Note that the central node lies closer to the global minimum than the corresponding part of the continuous system. Such configurations will become unstable when k is increased, since in the opposite, continuum limit there are only unstable saddle point configurations.

The diagram of stable configurations is presented in Fig. 6. The diagram has the following meaning: the one-node configuration (*i.e.*, with one node in the deeper well) is stable in the B-k below the 1st line, the two-node configuration is stable



Fig. 5. Stable system configuration corresponding to a bell-shape.

below the 2nd line, etc. Kinks are stable below the dashed line. Note that for any n, there is an area of parameters where only the n-node configuration is stable. For potentials closer to symmetric (lower B), there appear such areas for higher n, corresponding to more and more separated kinks.



Fig. 6. Stable configurations of a discrete system (see the explanation in the text). The inset shows the diagram for a wider range of *B*.

4. Final remarks

We have studied the properties of non-topological, bell-shaped excitations in the system with asymmetric potential, both in the continuum limit and in the discrete system. The bell-shaped configuration in the continuum limit is a saddle point of the potential energy. Due to the continuous translational symmetry, there is a family of equivalent saddle points. This property is manifested through the presence of the pseudo-Goldstone mode.

In the discrete system the potential energy around the configuration analogous to the bell-shape continuum limit becomes rough. For systems close to the continuum limit this roughness consists simply in replacing the original continuum of equivalent saddle points with isolated saddle points separated by "hills" (saddle points with two negative curvatures). This might seem analogous to the Peierls–Nabarro potential for kinks, but unlike the latter, the potential energy for bell-shaped waves cannot be defined rigorously, and is of little importance, since such waves are not stable. For strongly discrete systems, the shape of the potential energy becomes essentially different: many saddle points and local minima may appear. These characteristic points may be numerically searched and classified.

References

- [1] J. Krumhansl and J. Schrieffer. Phys. Rev. B, 11, 3535, (1975).
- [2] J. Currie, S. Trullinger, A. Bishop, and J. Krumhansl. Phys. Rev. B, 15, 5567, (1977).
- [3] V. Makhankov. Phys. Rep., 35, 1, (1978).
- [4] M. Collins, A. Blumen, J. Currie, and J. Ross. Phys. Rev. B, 19, 3630, (1979).
- [5] E. Kryachko. Solid State Commun., 65, 1609, (1988).
- [6] Y. Ishimori and T. Munakata. J. Phys. Soc. Japan, 51, 3367, (1982).
- [7] J. Combs and S. Yip. Phys. Rev. B, 28, 6873, (1983).
- [8] M. Peyrard and M. D. Kruskal. Physica D, 14, 88, (1984).
- [9] C. Willis, M. El-Batanouny, and P. Stancioff. Phys. Rev. B, 33, 1904, (1986).
- [10] P. Stancioff, C. Willis, M. El-Batanouny, and S. Burdick. Phys. Rev. B, 33, 1912, (1986).
- [11] R. Boesch, P. Stancioff, and C. Willis. Phys. Rev. B, 38, 6713, (1988).
- [12] R. Boesch and C. Willis. Phys. Rev. B, 39, 361, (1989).
- [13] R. Blinc and D. Hadzi. Mol. Phys., 1, 391, (1958).
- [14] J. Nagle, M. Mille, and H. Horowitz. J. Chem. Phys., 72, 3959, (1980).
- [15] Y. Kashimori, T. Kikuchi, and K. Nishimoto. J. Chem. Phys., 77, 1904, (1982).
- [16] A. Gordon. Physica B, 146, 373, (1987).
- [17] L. Manevich, V. Smirnov, A. Savin, and S. Volkov. Uspekhi Fiz. Nauk, 164, 937, (1994).
- [18] A. Gordon. Physica B, 150, 319, (1988).
- [19] A. Gordon. Solid State Commun., 72, 223, (1989).
- [20] A. Gordon. Z. Phys. B, 96, 517, (1995).
- [21] H. Metiu, K. Kitahara, and J. Ross. J. Chem. Phys., 64, 292, (1976).
- [22] J. Xu and J. Huang. Phys. Lett. A, 197, 127, (1995).
- [23] J. Xu and B. Zhou. Phys. Lett. A, 210, 307, (1996).
- [24] A. Gordon. J. Phys: Cond. Matt., 8, 4855, (1996).
- [25] S. Volkov and A. Savin. Ukr. Fiz. Zh., 37, 498, (1992).
- [26] S. Volkov. Phys. Lett. A, 224, 93, (1996).
- [27] P. Machnikowski and A. Radosz. Phys. Lett. A, 242, 313, (1998).
- [28] R. Rajaraman. Solitons and Instantons. North-Holland Publishing Company, Amsterdam, (1982).
- [29] R. Jackiw. Rev. Modern Phys., 49, 682, (1977).