

Huygens' Principle as Universal Model of Propagation



Peter Enders

Ahornallee 11, D-15754 Senzig, Germany.

E-mail: enders@dekasges.de

(Received 2 December 2008; accepted 12 January 2009)

Abstract

Huygens' Principle (HP) contains both the principle of action-at-proximity and the superposition principle. Although the propagation of sharp, non-spreading wave fronts is included in Huygens' (1690) original formulation, it can be left out without touching those principles. The formulation of HP by means of the Chapman-Kolmogorov equation (following Feynman 1948) comprises both versions and overcomes misunderstandings like "Huygens' principle is not exactly obeyed in Optics" (Feynman 1948) and "HP is incompatible with Green's functions" (Johns 1974). This way, HP applies not only to the propagation of light, but also to heat and matter diffusion, Schroedinger matter waves, ie, to virtually all propagation phenomena, which can be described through explicit linear differential and difference equations, respectively. HP for Maxwell's equations is discussed in terms of the Helmholtz-decomposed fields and currents. The appearances of HP in mechanics and in fractional Fourier transformation being exploited in optics are also mentioned.

Keyword: Optics, Huygens Principle, light propagation.

Resumen

El Principio de Huygens (PH) contiene tanto al principio de acción a proximidad como al principio de superposición. Aunque la propagación del pico no esparcido de los frentes de onda está incluido en el trabajo de Huygens (1690) acerca de la ecuación de Chapman-Kolmogorov (que sigue al trabajo de Feynman de 1948), incluye a ambas versiones y sobrepasa a los malentendidos como los de que "El principio de Huygens no es exactamente obedecido en óptica" (Feynman 1948) y "PH es incompatible con las funciones de Green" (Johns 1974). De esta forma, el PH se aplica no solamente en la propagación de la luz, pero también en la difusión del calor y la materia, en las ondas de Schrodinger de materia, es decir, a virtualmente toda la fenomenología de la propagación, la cual puede ser descrita a través de ecuaciones diferenciales lineales explícitas, respectivamente. El PH para las ecuaciones de Maxwell es discutido en términos de la descomposición de los campos y corrientes. En este trabajo mencionamos como es que la aparición del PH en mecánica y en las transformadas fraccionales de Fourier está siendo explotada en el campo de la óptica.

Palabras clave: Óptica principio de Huygens' propagación de la luz.

PACS: 11.10.-z, 41.85.-p, 42.15.-i, 42.15.Dp

ISSN 1870-9095

I. INTRODUCTION

No one doubts that physics is an exact science. Nevertheless, the notion 'exact science' "should not be interchanged with 'like mathematics'. As stressed by Huygens (1990, p. III f.), within physics, "one will find proofs of a kind, which do not grant the same great certainness of that of geometry and which even are rather different from those, because here, the principles are verified by the conclusions drawn from them, while the geometricians proof their theorems out of sure and unquestionable principles; the nature of the subjects dealt with conditions this".

Huygens' ideas on how light propagates have become basic ingredients of our physical picture of the world. The notion Huygens' principle (HP), however, is not uniquely used. This paper aims, on the one hand, at the clarification of some confusion existing in the literature, in particular, about the role of sharp, non-spreading wave fronts and the range of applicability. For instance, Feynman (1948) wrote, that HP holds exactly for wave mechanics, but only

approximately for optics, and Scharf (1994) stated, that HP is a principle of geometrical optics, not of wave optics. On the contrary, the unifying power of HP will be demonstrated here.

Some of that confusion is related to Kirchhoff's formula and reaches up to doubts on the validity of HP at all (Miller 1991), or on the possibility of the representation of HP by means of Green's functions (GF) (Johns 1974). Both doubts contradict any mind believing in the unity of physics. Indeed, Kirchhoff's solution to the wave equation, while being mathematically exact, suffers from the drawback of requiring the knowledge of both the field amplitude and its gradient on the boundaries. I will trace the origin of these mathematical and physical difficulties to the notions of degrees of freedom of motion and of independent dynamical variables.

For the sake of the unity of physics, a further goal of this paper is to generalize Huygens' basic ideas. This means, that I will keep essentially the imagination, that each locus of a wave excites the local matter which reradiates a secondary wavelets, and all wavelets

superpose to a new, resulting wave (the envelope of those wavelets), and so on. Huygens' *ad-hoc* omission of backward radiation as well as Fresnel's and other auxiliary assumptions (cf.

Longhurst 1973, §10-2) is requested to be included in a natural manner. In particular, attention will be paid to a simple, but general and exact description of wave and other propagation processes, which obey the principle of action-by-proximity and can be described by explicit transport equations.

Shortly, consider a complete set of independent dynamical variables of a given problem, $\vec{X}(\vec{r}, t) = (X_1(\vec{r}, t), \dots, X_f(\vec{r}, t))$, eg, $\vec{X}(\vec{r}, t) = (u(\vec{r}, t), \partial u(\vec{r}, t)/\partial t)$, $u(\vec{r}, t)$ being the amplitude of a scalar wave. I seek to represent its propagation in the most simple form

$$\vec{X}(\vec{r}, t) = \iiint \hat{H}(\vec{r}, t; \vec{r}_0, t_0) \cdot \vec{X}(\vec{r}_0, t_0) d^3\vec{r}_0; \quad t > t_0 \quad (1)$$

The "Huygens propagator", \hat{H} , obviously, obeys the CHAPMAN-KOLMOGOROV equation (KOLMOGOROV 1931, 1933, CHAMPMAN & COWLING 1939) known from (but not restricted to) MARKOV processes and related problems of probability theory.

$$\hat{H}(\vec{r}, t; \vec{r}_0, t_0) = \iiint \hat{H}(\vec{r}, t; \vec{r}_1, t_1) \cdot \hat{H}(\vec{r}_1, t_1; \vec{r}_0, t_0) d^3\vec{r}_0; \quad t > t_1 > t_0 \quad (2)$$

Thus, following Feynman (1948), I will express HP through this equation. The rigorous treatment requires measure theory (Dynkin 1965), but this is much more than necessary for the understanding of 'common' physical propagation processes. It may proven useful, however, for the fractal description of wave propagation in disordered media (West 1992) and the like.

If $\vec{X}(\vec{r}, t)$ obeys a set of partial differential equation of first order in time, $\hat{H}(\vec{r}, t; \vec{r}_0, t_0)$ turns out to be the GF of that equation, and Eq.(1) is the solution to the initial-boundary value problem. If, however, $\vec{X}(\vec{r}, t)$ obeys a set of partial differential equation of second (or higher) order in time, no such simple equation exists. Often, the much more involved Kirchhoff's formula (11) is used. This has misled some authors to deny a relationship between GF, HP and wave propagation at all.

The use of GF within such considerations is not new, of course (Courant et al 1928, Spitzer 1964, Keilson 1965). However, our goal is the representation of HP through GF rather than a discussion of the probabilistic questions behind such approaches. These are interesting enough, but need (and deserve!) a separate treatment. We will encounter discrete Markov processes when discussing computational algorithms realizing HP in discrete form.

Such forms are required for numerical calculations on digital computers. The natural formulation is in terms of Markov chains. On the basis of transmission-line networks, powerful algorithms have been developed not only for electromagnetic problems, but also for diffusion and even for mechanical problems (Hoefer & So 1991,

Christopoulos 1995, de Cogan 1998, de Cogan et al 2005). Because here – in contrast to other cellular automata algorithms (Chopard & Droz 1998) –, an (idealized) physical system is mapped, it is not too surprising that HP applies to the TLM equations, too (Hoefer 1991, Enders 2001, Enders & Vanneste 2003). Therefore, some implications of our approach to HP for practical, in particular, wave-optical computations will also be discussed.

For historical and methodological reasons, I start in Section 2 with HP in mechanics and continue, in Section 3, with Kirchhoff's formula and certain problems of its physical interpretation. Then, Hadamard's rigorous definition of HP is discussed. In section 5, the superposition of secondary wave (let)s is represented and illustrated by means of general field propagators in the space-time domain. This leads to a description of wave motion, that overcomes the difficulties in the interpretation and application of Kirchhoff's formula mentioned above. Section 6 stresses the role of time-derivatives of dynamical variables as independent dynamical variables. When equations of 2nd order in time, such as the wave equation, are rewritten as systems of equations of 1st order in time, HP applies exactly to those and, consequently, to wave optics as well. Section 7 discusses Maxwell's equations in the light of these results, where the fields and currents are Helmholtz-decomposed, in order to work with independent field variables only. Section 8 applies these thoughts to difference equations and discusses implications for practical computations. A relationship to the fractional Fourier transformation is sketched in section 9. Section 10 condenses these results into thesis for the general formulation of the physics of propagation. Sections 11, finally, summarizes and concludes the results.

II. HUYGEN'S PRINCIPLE IN MECHANICS

A.Principle for the free fall

As a matter of fact, the principle of superposition has first been formulated by Huygens for mechanical motions. Shortly, during free fall, the momentarily achieved increments of speed add to the speed assumed just before (Horologium oscillatorium, 1673; after Simonyi, 1990, p.241f.). This implies the differentiability of the velocity: $\vec{v}(t + dt) = \vec{v}(t) + d\vec{v}$, therefore, the smoothness of the trajectories.

B.Huygens' construction for the classical harmonic oscillator

The trajectory, $x(t)$, of an 1D harmonic oscillator can be described as function of the initial values of location, $x(0)$, and momentum, $p(0)$, and of its mass, m , and angular velocity, w .

$$x(t) = x(0)\cos(wt) + \frac{p(0)}{mw}\sin(wt) \quad (3)$$

Here, the internal (m, w) and external parameters $(x(0), v(0))$ occur in mixed form. Since, generally speaking, separations highlight the actual physical interrelations, it is desirable to separate internal und external parameters, *ie*, in the case, the constants (laws of motion, system parameters) from the variable influences (initial conditions), *cf* (WIGNER 1963).

The separation makes it immediately, if one writes down the coupled solutions for both dynamical variables, $x(t)$ and $p(t)$:

$$\begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = \begin{pmatrix} \cos(wt) & \frac{1}{Mw} \sin(wt) \\ -Mw \sin(wt) & \cos(wt) \end{pmatrix} \begin{pmatrix} x(0) \\ p(0) \end{pmatrix} \stackrel{\text{def}}{=} \widehat{D}(t) \begin{pmatrix} x(0) \\ p(0) \end{pmatrix}$$

This form emerges, when one solves Hamilton's equations of motion as a *system* of equation. The (matrix-valued) propagator $\widehat{D}(t)$ contains solely the internal parameters and the time. It describes rotations in phase space $\{x(t), p(t)\}$ and exhibits the group property

$$\widehat{D}(t) = \widehat{D}(t-t') \cdot \widehat{D}(t'); \quad 0 \leq t' \leq t \quad (4)$$

This is an example for HUYGENS' construction and Eq. (4) a discrete form of the CHAPMAN-KOLMOGOROV equation.

Finally, one should separate angular frequency and mass; for the oscillation is determined by the former only. This is possible through the diagonalization of $\widehat{D}(t)$ (what else?):

$$\begin{pmatrix} \tilde{x}(t) \\ \tilde{p}(t) \end{pmatrix} = \begin{pmatrix} e^{iwt} & 0 \\ 0 & e^{-iwt} \end{pmatrix} \begin{pmatrix} \tilde{x}(0) \\ \tilde{p}(0) \end{pmatrix} = \widehat{\widehat{D}}(t) \begin{pmatrix} \tilde{x}(0) \\ \tilde{p}(0) \end{pmatrix}. \quad (5)$$

The transformed variables are

$$\begin{pmatrix} \tilde{x}(t) \\ \tilde{p}(t) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ iMw & 1 \end{pmatrix} \begin{pmatrix} x(t) \\ p(t) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} x(t) + \frac{i}{Mw} p(t) \\ iMw x(t) + p(t) \end{pmatrix}. \quad (6)$$

They exhibit the most simple time dependences

$$\tilde{x}(t) = \tilde{x}(0)e^{iwt}; \quad \tilde{p}(t) = \tilde{p}(0)e^{-iwt}.$$

From them, two first integrals of motion can be read off immediately,

$$I_1 = e^{-iwt} \tilde{x}(t) = \tilde{x}(0); \quad I_2 = e^{iwt} \tilde{p}(t) = \tilde{p}(0).$$

Since there are no further independent first integrals, the total energy can be represented as a function of I_1 and I_2 . Indeed, $E = -i w I_1 I_2$. As a consequence, the variables (6) *factorize* the Hamilton function (??).

$$\begin{aligned} H(x, p) &= \frac{M}{2} w^2 x^2(t) + \frac{1}{2M} p^2(t) = -i w \tilde{x}(t) \tilde{p}(t), \\ &= -i w \tilde{x}(0) \tilde{p}(0) = E. \end{aligned}$$

Finally, the variables (6) obey equations motion no t of second order, as $x(t)$ and $p(t)$, but of first order in time.

$$\frac{d}{dt} \tilde{x}(t) = +i w \tilde{x}(t); \quad \frac{d}{dt} \tilde{p}(t) = -i w \tilde{p}(t).$$

The factorization of that equation is obvious, as $\left(\frac{d^2}{dt^2} + w^2\right) = \left(\frac{d}{dt} + i w\right) \left(\frac{d}{dt} - i w\right)$.

It is noteworthy that this result was possible only by means of the imaginary unit, $i \equiv \sqrt{-1}$. This provides *i* with a *physical* (and not only mathematical-calculational) justification already within classical mechanics (SCHROEDINGER 1926 hesitated to exploit *i* for the formulation of the first-order time-dependent SCHROEDINGER equation).

C. Huygens' principle in quantum mechanics

Being a probabilistic theory, quantum mechanics describes motion in terms of transition probabilities, $P_{ba} = \langle b | a \rangle$. Rather than trajectories from state *a* to state *b*. Usually, these states form a complete set such, that they provide a decomposition of the unit operator, $\hat{1} = \sum_c |c\rangle \langle c|$. This enables us to write

$$P_{ba} = \langle b | a \rangle \geq \sum_c \langle b | c \rangle \langle c | a \rangle \geq \sum_c P_{bc} P_{ca}. \quad (7)$$

Referring to this equation, FEYNMAN (1948) emphasized, that HP applies to SCROEDINGER Wave Mechanics. Therefore, FEYNMAN considered the CHAPMAN-KOLMOGOROV equation (here, in state space) to be the mathematical expression of HP.

III. KIRCHHOFF'S FORMULA AND ITS DIFFICULTIES OF INTERPRETATION AND APPLICATION

Within classical wave theory, the mathematical problem of wave propagation is usually reduced to the solution of the wave equation, *ie* in the simplest 3D case,

$$\square u(\vec{r}, t) \equiv \Delta u(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 u(\vec{r}, t)}{\partial t^2} = -q(\vec{r}, t). \quad (8)$$

Where $u(\vec{r}, t)$ is the scalar field amplitude and $q(\vec{r}, t)$ the source density. The GF, $g(\vec{r}, t | \vec{r}_0, t_0)$, to Eq. (8) is its solution for the unit source density,

$$\square g(\vec{r}, t | \vec{r}_0, t_0) = -\delta(\vec{r} - \vec{r}_0) \delta(t - t_0). \quad (9)$$

A special solution to Eq. (9) is the expanding (retarded=impulsive spherical wave,

$$g_r(\vec{r}, t | \vec{r}_0, t_0) = \frac{\theta(\tau) \delta\left(\frac{R}{c} - \tau\right)}{4\pi R}; \quad R \equiv |\vec{r} - \vec{r}_0|; \quad \tau \equiv t - t_0. \quad (10)$$

(DESANTO 1992, p. 47) Using this GF, the retarded solutions to Eq. (8) read

$$u(\vec{r}, t) = \int_{t_0}^t dt_0 \iiint g_r(\vec{r}, t | \vec{r}_0, t_0) q(\vec{r}_0, t_0) dV_0$$

$$\begin{aligned}
 & -\frac{1}{c^2} \iiint \left[u(\vec{r}_0, t_0) \frac{\partial}{\partial t} g_r(\vec{r}, t | \vec{r}_0, t_0) \frac{\partial}{\partial t} u(\vec{r}_0, t_0) \right] dV_0 \\
 & + \int_{t_0}^t dt_0 \iint_{\partial V_0} \left[g_r(\vec{r}, t | \vec{r}_0, t_0) \nabla_0 u(\vec{r}_0, t_0) - \right. \\
 & \left. u(\vec{r}_0, t_0) \nabla_0 g_r(\vec{r}, t | \vec{r}_0, t_0) \right] \cdot d\vec{S}_0. \quad (11)
 \end{aligned}$$

The first term describes the propagation of that part of the field amplitude that stems from the external sources, while the second and third terms account for the initial and boundary conditions, respectively.

In particular, the third term, KIRCHHOFF'S formula, describes the scattering at screens. Its physical interpretation is subject to various discussions. MACKE (1958) designates this term as a "direct and quantitative formulation of HUYGENS' principle" (cf. also NAAS & SCHMID 1974). On the other hand, KIRCHHOFF'S formula has been criticized as not only the field amplitude itself (which may be complex-valued) is called in, but also its spatial derivative, and that derivations has "to be performed by nature" (JOHNS 1974, MILLER 1991). Thus, "two types of sources of varying strength occur, so that the simplicity of HUYGENS' approach is lost" (MILLER 1991).

It seems to us, however, that these problems originates not from HP, but from, (i) the use of the *free-space* GF (10), which does *not* account for the *actual* boundary conditions, such as screens, and, (ii), from the use of a GF which does not represent HP by itself. Indeed, they disappear when rewriting the wave equation (8) as

$$\frac{\partial u(\vec{r}, t)}{\partial t} = w(\vec{r}, t), \quad (12)$$

$$\frac{\partial w(\vec{r}, t)}{\partial t} = c^2 \Delta u(\vec{r}, t) + c^2 q(\vec{r}, t), \quad (13)$$

and using the HUYGENS propagator for this system of equations (for more details, see below).

IV. HADAMARD'S NOTION OF HUYGENS' PRINCIPLE

Basing on his investigations on the CAUCHY problem for partial differential equations, HADAMARD (1952, § 33) has given the most exact formulation of HP I am aware of.

A. Hadamard's syllogism

(A) syllogism is a form of logical conclusion, which has been developed already by ARISTOTLE (LUKASIEWICZ, 1957, WIENER 1995). The conclusion is derived from two premises, a major and a minor ones.

(B) **Minor premise.** The propagation of light pulses proceeds without deformation (spreading, tail building) of the pulse;

(C) **Conclusion.** In order to calculate the effect of our initial luminous phenomenon produced at $t = t_0$, one may

replace it by a proper system of disturbances taking place at $t = t'$ and being distributed over the surface of the sphere with centre t_0 and radius $c(t' - t_0)$.

The Major premise (A) is the principle of action-by-proximity and, philosophically speaking, a "law of thought". The Minor premise (B) postulates the propagation of none-spreading wave fronts. The conclusion (C) is essentially HUYGENS' construction.

As a matter of fact, in the Conclusion (C), the isotropy of re-irradiation can be replaced with the re-irradiation according to the actual local propagation conditions. This means, that the secondary sources represent the local propagation properties of the material under consideration (or that of free space). For instance, in anisotropic media, the reaction of the secondary sources is anisotropic, while in nonlinear media, their excitation and re-irradiation is not proportional to the amplitude of the exciting field (GUENTHER 1988).

Now, when compared with Major premise, the Minor premise is rather special. It is necessary for geometrically constructing the wave front, but not for the basic principle of action-by-proximity and not for the cycle of excitation and re-irradiation. One of the main points of this paper is to examine what happens, when it is removed.

B. 'Special' Huygens' principle: Minor Premise included

Often, the existence of sharp, non-spreading wave fronts is already referred to as HP (COURANT & HILBERT 1968, NAAS & SCHMIDT 1974, <http://www.mathpages.com/home/kmath242/kmath242.htm>). This phenomena is widely known for D'ALEMBERT'S wave equation. Another classical example is the distortion-free pulse propagation along special 1D transmission lines (HEAVISIDE, PUPIN). HADAMARD'S conjecture states that the wave front is not spreading in odd space dimensions (HADAMARD 1953, NAAS & SCHMIDT 1974). It should be clear, however, that, despite of its practical consequences for signal transmission, the Minor premise (B) is a secondary attribute of propagation processes, while the Major premise (A) and the Conclusion (C) are primary ones.

Thus, analogously to the solution of total-hyperbolic differential equations (NAAS & SCHMIDT 1974), one may define also within optics and for general propagation phenomena, respectively, a 'Special' HUYGENS' Principle, where proposition (B) is an essential ingredient. It corresponds to a distortion-free signal transmission; the speed of propagation of the waves does not depend on the oscillation frequency of the source and the waves suffer not any deformation through smearing or wake building (cf also IWANENKO & SOKOLOV 1953). In this case, for the validity of HP it is necessary and sufficient that Green's function of D'ALEMBERT'S wave equation is proportional to the delta-function $\delta(R - c\tau)$ or to its derivatives (NAAS & SCHMIDT 1974).

The construction of equations the solutions to which are non-spreading sharp wave fronts has been developed to

a special topic of its own (see, *eg*, GUENTHER 1988). These results may proven to be useful for the design of dispersionless signal transmission systems. Is there a relationship to reflectionless potentials for the SCHRODINGER equation?

C. 'General' Huygens' principle: Minor Premise *not* included

On the other hand, action-at-proximity and superposition are not bounded to sharp wave fronts neither is HUYGENS' construction, as we will show below. Thus, Major premise (A) together with Conclusion (C) has also be termed HP, *eg*, by JOHNS (1974), BICKEL & HANDER (1988), MILLER (1991). We will call HP the combination of action-at-proximity ("elastic waves in aether" in HUYGENS' pictorial imagination) and superposition of secondary wavelets (HUYGENS' construction, suitably generalized). The shape of the wave front may vary from case to case, without influence on these basic ingredients of propagation, but the essentials of HUYGENS' (and Faraday's) imagination of propagation are conserved. The advance of this notion of HP consists in that its applicability becomes extremely wide; in fact, in this form, HP qualifies to a clue for unifying the physical and mathematical description of a huge variety of transport and propagation processes.

V. GREEN'S FUNCTIONS FOR REPRESENTING HUYGENS' PRINCIPLE

From the theoretician's point of view, GF represent one of the most powerful and, at once, most beautiful and clear (propagator!) tools of mathematical physics at all (DYSON 1993). Therefore, it is naturally expected that there are GF which do provide a representation of HP.

A. Huygens propagators

Without loss in generality, let us study the equation

$$u(\vec{r}, t) = \iiint H(\vec{r}, t; \vec{r}', t') u(\vec{r}', t') d^3\vec{r}'; \quad t > t', \quad (14)$$

describing the propagation of the scalar field $u(\vec{r}, t)$ from the space-time point (\vec{r}', t') to the space-time point (\vec{r}, t) . Which are the general properties of the integral kernel, H ?

i. If $u(\vec{r}, t)$ fulfills the partial differential equation

$$\frac{\partial u}{\partial t} = L(\vec{r})u(\vec{r}, t) - q(\vec{r}, t), \quad (15)$$

where $L(\vec{r})$ is a partial differential expression in \vec{r} and $q(\vec{r}, t)$ the source density, then

Huygens' Principle as Universal Model of Propagation

$$\begin{aligned} \frac{\partial}{\partial t} H(r, t; r_0, t_0) &= \\ &= L(\vec{r})H(r, t; r_0, t_0) - \delta(\vec{r} - \vec{r}_0)\delta(t - t_0). \end{aligned} \quad (16)$$

Thus, H is the Green's function of the differential equation (15). If $u(\vec{r}, t)$ fulfills a partial differential equation of *higher* order in time, there is no simple relationship (14).

ii) $u(\vec{r}, t)$ fulfills the initial condition

$$u(\vec{r}, t=0) = u_0(\vec{r}). \quad (17)$$

If H obeys the initial condition

$$\lim_{t \rightarrow t_0+0} H(r, t; r_0, t_0) = \delta(\vec{r} - \vec{r}_0), \quad (18)$$

and

$$\begin{aligned} u(\vec{r}, t) &= \iiint H(\vec{r}, t; \vec{r}', 0) u_0(\vec{r}') d^3\vec{r}', \\ t &> 0, \end{aligned} \quad (19)$$

iii) $u(\vec{r}, t)$ fulfills the boundary condition

$$B(\vec{r})u(\vec{r}, t) = 0; \quad \vec{r} \in S, \quad (20)$$

on the inner surface, S , of the domain considered, if H does so,

$$B(\vec{r}')H(\vec{r}, t; \vec{r}_0, t_0) = 0; \quad \vec{r} \in S, \vec{r}' \notin S. \quad (20)$$

This can be achieved by means of an eigenfunction expansion (DeSanto 1992).

Thus, a Huygens *propagator*, H , is a GF (16) of a differential equation of first order in time (15) which, additionally, obeys the initial condition (18) and the boundary conditions (21) of the problem under consideration. In other words, it contains both the propagation conditions in the volume *and* on the spatio-temporal boundaries. Due to this, the difficulties with the boundary terms in Kirchhoff's formula are overcome.

B. The Chapman-Kolmogorov equation as generalization of Huygens' construction

Nesting the integral equation (1) into itself yields

$$\begin{aligned} u(\vec{r}, t) &= \iiint H(\vec{r}, t; \vec{r}_0, t_0) \times \\ &\quad \iiint H(\vec{r}_0, t_0; \vec{r}_1, t_1) u(\vec{r}_1, t_1) d^3\vec{r}_1 d^3\vec{r}_0; \\ t &> t_0 > t_1 \end{aligned} \quad (22)$$

Rearranging the indices and comparing this with the original equation (1) gives

$$H(\vec{r}, t; \vec{r}_0, t_0) = \iiint_{V_1} H(\vec{r}, t; \vec{r}_1, t_1) \times H(\vec{r}_1, t_1; \vec{r}_0, t_0) d^3\vec{r}_1; \quad (23)$$

$$t > t_0 > t_1.$$

This is the Chapman-Kolmogorov equation in the space-time domain. It generalizes Huygens' construction to spreading wave fronts; the domain of sources of secondary wavelets is not necessarily a surface, but, in general, a certain volume, V_1 . Non-spreading wave fronts correspond to δ -functions in the GF reducing the volume integral to a surface integral. In such cases, diffraction at screens is treated in a manner resembling Kirchhoff's formula, but without its difficulties mentioned above.

Since the time interval $t - t_0$ can be infinitesimally small, the Chapman-Kolmogorov equation is a mathematical formulation not only of the superposition of secondary wavelets, but also of the action-at-proximity.

The validity of a relation like (23) is sometimes called a Markov property; it plays an important role for the path-integral representation of dynamical processes (Feynman & Hibbs 1965).

Now, as a matter of fact, the GF (10) of the wave equation (8) does *not* obey the Chapman-Kolmogorov equation (23). Indeed, the latter is obeyed by functions being the solution to partial differential equations of *first* order in time (this may be easily proven by means of the Fourier transformation w.r.t. the time variable). This was, perhaps, the reason for Feynman (1948) to state that, in optics, HP holds true only approximately.

C. Treatment of differential equations of higher-order in time

The way out consists in that, for wave and other propagation processes of higher order in time, one has to 'return' to systems of first-order equations. Remarkably enough, these are often the *fundamental* relations, *viz.*, constitutive equation(s) and conservation law(s).

For instance, the Maxwell equations supplemented with appropriate constitutive equations connecting the field strengths with the inductions represent such a system. Another example is the following system of equations of hyperbolic heat conduction theory (Mueller 1967),

$$\frac{\partial}{\partial t} \begin{pmatrix} \vec{J}(r, t) \\ \vec{T}(r, t) \end{pmatrix} = - \begin{pmatrix} \frac{1}{\tau} & \frac{\lambda}{\tau} \nabla \\ \frac{1}{\rho C_p} \nabla \cdot & 0 \end{pmatrix} \begin{pmatrix} \vec{J}(r, t) \\ \vec{T}(r, t) \end{pmatrix}, \quad (24)$$

(J heat current density, T temperature, τ heat flux relaxation time, λ heat conductivity, ρC_p heat capacity per unit volume at constant pressure). The physics behind the necessity to work with first-order equations will be discussed in the next section.

The GF for the *system* of equations (24) is the 4×4 *matrix*-valued function

$$\hat{G} = \begin{pmatrix} \hat{G}^{\vec{J}\vec{J}} & \hat{G}^{\vec{J}\vec{T}} \\ \hat{G}^{\vec{T}\vec{J}} & \hat{G}^{\vec{T}\vec{T}} \end{pmatrix}, \quad (25)$$

being defined as the solution to the *matrix*-valued equation

$$\frac{\partial \hat{G}}{\partial t} = \begin{pmatrix} \frac{1}{\tau} & \frac{\lambda}{\tau} \nabla \\ \frac{1}{\rho C_p} \nabla \cdot & 0 \end{pmatrix} \cdot \hat{G} + \begin{pmatrix} \hat{1} & \vec{0} \\ \vec{0} & \hat{1} \end{pmatrix} \delta(\vec{r} - \vec{r}') \delta(t - t'). \quad (26)$$

When $\hat{G}(\vec{r}, t; \vec{r}', t')$ also accounts for the actual initial $\vec{J}_0(\vec{r}) = \vec{J}(\vec{r}, t = 0)$, $\vec{T}_0(\vec{r}) = \vec{T}(\vec{r}, t = 0)$ and boundary conditions, it becomes the Huygens propagator, $\hat{H}(\vec{r}, t; \vec{r}_0, t_0)$, of the considered problem, and the solution (\hat{J}, T) is given by a *single* integral, again.

$$\begin{pmatrix} \vec{J}(r, t) \\ \vec{T}(r, t) \end{pmatrix} = \iiint \begin{pmatrix} \hat{H}^{\vec{J}\vec{J}} & \hat{H}^{\vec{J}\vec{T}} \\ \hat{H}^{\vec{T}\vec{J}} & \hat{H}^{\vec{T}\vec{T}} \end{pmatrix} \begin{pmatrix} \vec{J}_0(r, t) \\ \vec{T}_0(r, t) \end{pmatrix} d^3\vec{r}; \quad (24)$$

$$t > 0$$

VI. THE TIME DERIVATIVE OF THE WAVE AMPLITUDE AS INDEPENDENT DYNAMICAL VARIABLE

Huygens propagators of the wave equation in the form (12) describe the *common* propagation of the field (wave) amplitude, $u(\vec{r}, t)$, and of its time-derivative,

$\frac{\partial}{\partial t} u(\vec{r}, t) = w(\vec{r}, t)$, as *independent dynamical variables*, which are *created simultaneously* and propagates *together and in mutual interaction*. When considering the pair (\vec{J}, T) of Eq. (24) or counter-propagating waves (d'Alembert's solution), the physical content of the derivative is even more obvious. This is the fundamental difference between our interpretation of HP and previous ones, but Hadamard (1952).

Nowadays mechanical theories often concentrate on equations of motion, such as Eq.(8), or Lagrange's equation of motion, while the role of velocity, $\vec{v} = d\vec{r}/dt$, and Cartesian momentum, $\vec{p} = m d\vec{r}/dt$, respectively, as dynamical variables on its own is explicitly considered only in Hamilton's equations of motion (and in statistical mechanics). However, the independence of the initial values of location, $\vec{r}_0 = \vec{r}(t = 0)$, and of velocity, $\vec{v}_0 = \vec{v}(0)$, implies the independence of the values of $\vec{r}(t)$

and $\vec{v}(t)$ for all later times. And this qualifies $\vec{r}(t)$ and $\vec{v}(t)$ as independent dynamical variables.

Obviously, the same holds true for (possibly complex-valued) wave amplitudes, ψ , and their time-derivates $\partial\psi/\partial t$, as recognized within Lagrangean and Hamiltonian field theories. Within optics, this matter of fact is commonly hidden by the use of time-harmonic waves. The Schroedinger equation is *one* common parabolic equation for the *two* independent dynamical variables, ψ and ψ^* , or $Re(\psi)$ and $Im(\psi)$.

We generalize these results in the following

Conjecture 1 *The number of independent dynamical variables is equal to the number of time-derivatives in the equation(s) of motion.*

In general, there are *various* complete sets of independent dynamical variables for a given problem.

His number can be reduced by symmetry. For instance, in a travelling electromagnetic wave in free space, all 12 field components in Maxwell's macroscopic equations are proportional to only *two* field components (*eg*, E_x and E_y) determining intensity and polarization.

Scalar propagators obeying the Chapman-Kolmogorov equation are positive definite. Hence, the Huygens propagators for processes exhibiting interference are matrices (classical waves) or complex-valued (matter waves).

The Hamilton-Jacobi equation (wave picture) converts the Hamiltonian equations of motion (particle picture) into a non-linear 'wave' equation (*cf* also Einstein 1917). It would thus be interesting to explore the applicability of HP to the former.

VII. MAXWELL'S EQUATIONS AND HUYGENS' PRINCIPLE

From the point of view of initial-boundary value problems, the Maxwell equations represent an incomplete set of partial differential equations of 1st order in time for the 12 field variables \vec{D} , \vec{B} , \vec{E} and \vec{H} .

$$\nabla \cdot \vec{D} = \rho, \quad (28a)$$

$$\nabla \cdot \vec{B} = 0, \quad (28b)$$

$$\frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E}, \quad (28c)$$

$$\frac{\partial \vec{E}}{\partial t} = -\nabla \times \vec{H}. \quad (28d)$$

We complement them through the simplified material equations

$$\tau_{\vec{E}} \frac{\partial \vec{B}}{\partial t} + \vec{E} = \frac{\vec{D}}{\epsilon_r \epsilon_0}, \quad (29a)$$

$$\tau_{\vec{H}} \frac{\partial \vec{H}}{\partial t} + \vec{H} = \frac{\vec{B}}{\mu_r \mu_0}, \quad (29b)$$

accounting for finite relaxation times ($\tau_{\vec{E}}, \tau_{\vec{H}}$). The source equations (28a, 28b) make the vectors \vec{D} and \vec{B} *not* to represent *three* independent dynamical variables each. This deficiency of 2 independent dynamical variables is usually ascribed to charge and energy conservation implicitly imposed. Why, then, the conservation of momentum and angular momentum do not diminish the number of independent dynamical variables? – As a consequence, the Huygens propagator is degenerated, and so-called spurious modes may appear in numerical calculations.

Now, in terms of the Helmholtz decomposition (Helmholtz 1858, Keller 2005), $\vec{B} = \vec{B}_T + \vec{B}_L$, is purely transverse: $\vec{B} = \vec{B}_T, \vec{B}_L = 0$, and only the transverse components of \vec{H} and \vec{E} enter the Maxwell equations. In turn, the charge conservation is related to solely the longitudinal components of \vec{D} and \vec{j} .

$$\nabla \cdot \vec{j}_L + \frac{\partial \rho}{\partial t} = 0; \quad \nabla \cdot \vec{D} = \rho. \quad (30)$$

For this, we can rewrite Eqs. (31) as

$$\nabla \cdot \vec{D}_L = \rho, \quad (31a)$$

$$\nabla \cdot \vec{B}_L = 0, \quad (31b)$$

$$\frac{\partial \vec{B}_T}{\partial t} = -\nabla \times \vec{E}_T, \quad (31c)$$

$$\frac{\partial \vec{D}_T}{\partial t} = \nabla \times \vec{H}_T - \vec{j}_T. \quad (31d)$$

It is seen that the Helmholtz decomposition genuinely relates the propagation of electromagnetic waves with the *transverse* field components only. Its drawback – and, perhaps, reason of low acceptance – consists in the fact that it is not Lorentz covariant, so that it has to be separately performed in each system of reference. The criterion of being compatible with special relativity is, however, not the Lorentz covariance, but the compatibility with the Poincare group (Dirac 1949).

Eqs. (29) reveal that it depends on the properties of the matter in which the electromagnetic field under consideration exists, how many independent dynamical variables are represented by \vec{E} and \vec{H} . Without loss of generality, I confine myself to the simplest case, *viz*, that of vacuum, $\vec{E} = \vec{D}/\epsilon_0, \vec{H} = \vec{B}/\mu_0$. Note, that in contrast to the common use to work with the pair (\vec{E}, \vec{H}) , here, I keep the pair (\vec{D}, \vec{B}) in view of its position in the Maxwell equations, and not to contradict Mie (1941) and Sommerfeld (2001).

Let us further assume that

$$\vec{D}_T = \begin{pmatrix} D_x(z) \\ D_y(z) \\ 0 \end{pmatrix}; \vec{B}_T = \begin{pmatrix} B_x(z) \\ B_y(z) \\ 0 \end{pmatrix}; \vec{j}_T = \begin{pmatrix} j_x(z) \\ j_y(z) \\ 0 \end{pmatrix}. \quad (32)$$

Then, the four independent dynamical variables $(B_{x,y}, D_{x,y})$ obey the *complete* set of equations

$$\frac{\partial}{\partial t} \begin{pmatrix} B_x \\ B_y \\ D_x \\ D_y \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{\varepsilon_0} \frac{\partial}{\partial z} \\ 0 & 0 & -\frac{1}{\varepsilon_0} \frac{\partial}{\partial z} & 0 \\ 0 & -\frac{1}{\mu_0} \frac{\partial}{\partial z} & 0 & 0 \\ \frac{1}{\mu_0} \frac{\partial}{\partial z} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} B_x \\ B_y \\ D_x \\ D_y \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ j_x \\ j_y \end{pmatrix}. \quad (33)$$

The corresponding GF,

$$\hat{G} = \begin{pmatrix} G^{B_x B_x} & G^{B_x B_y} & G^{B_x D_x} & G^{B_x D_y} \\ G^{B_y B_x} & G^{B_y B_y} & G^{B_y D_x} & G^{B_y D_y} \\ G^{D_x B_x} & G^{D_x B_y} & G^{D_x D_x} & G^{D_x D_y} \\ G^{D_y B_x} & G^{D_y B_y} & G^{D_y D_x} & G^{D_y D_y} \end{pmatrix}, \quad (34)$$

obeys the same equation, where the vector $(0 \ 0 \ j_x \ j_y)$ is replaced with $\hat{1} \delta(z - z') \delta(t - t')$, $\hat{1}$ denoting the 4×4 -matrix GF is *reducible* as it is the direct product of the 2×2 -matrix GF of the two equations, into which Eq.(33) separates.

$$\frac{\partial}{\partial t} \begin{pmatrix} B_x \\ D_y \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{\varepsilon_0} \frac{\partial}{\partial z} \\ \frac{1}{\mu_0} \frac{\partial}{\partial z} & 0 \end{pmatrix} \begin{pmatrix} B_x \\ D_y \end{pmatrix} - \begin{pmatrix} 0 \\ j_y \end{pmatrix}, \quad (35)$$

$$\frac{\partial}{\partial t} \begin{pmatrix} B_y \\ D_x \end{pmatrix} = \begin{pmatrix} 0 & -\frac{1}{\varepsilon_0} \frac{\partial}{\partial z} \\ -\frac{1}{\mu_0} \frac{\partial}{\partial z} & 0 \end{pmatrix} \begin{pmatrix} B_y \\ D_x \end{pmatrix} - \begin{pmatrix} 0 \\ j_x \end{pmatrix}. \quad (36)$$

The Fourier components $e^{i(\omega t - kz)}$ of the corresponding 2×2 -matrix GF obey the equations

$$\begin{pmatrix} i\omega & \pm \frac{ik}{\varepsilon_0} \\ \pm \frac{ik}{\mu_0} & i\omega \end{pmatrix} \begin{pmatrix} G_{11}^\pm & G_{12}^\pm \\ G_{21}^\pm & G_{22}^\pm \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \quad (37)$$

These are *irreducible*, but can be diagonalized like Eq.(35),

$$\frac{\partial^2}{\partial t^2} D_y - \frac{1}{\varepsilon_0 \mu_0} \frac{\partial^2}{\partial z^2} D_y = \frac{\partial}{\partial t} j_y. \quad (38)$$

As in the 3D case, it is straightforward to show that the GF of the 1D wave equation (38) does not obey the Chapman-Kolmogorov equation, while the 2×2 -matrix GF of Eq. (35) does.

VIII. DISCRETE MODELS OF PROPAGATION

Let us illustrate these thoughts by means of simple discrete propagation models.

A. One-step Markov chains-discrete Huygens propagators

In one-step Markov chains, each two subsequent states, $\vec{u}_k \equiv (u_{k,1}, u_{k,2}, \dots)$ and \vec{u}_{k+1} , where the second index may label spatial cells, are connected through a transition matrix, \hat{P} , as

$$\vec{u}_{k+1} = \hat{P} \cdot \vec{u}_k; k = 0, 1, \dots \quad (39)$$

For k -independent \hat{P} , one obtains

$$\vec{u}_k = \hat{P}^k \cdot \vec{u}_0 \quad (40)$$

$$= \hat{P}^{k-1} \cdot \vec{u}_1 = \hat{P}^{k-1} \cdot (\hat{P} \cdot \vec{u}_0) \quad (41)$$

From this, the fundamental formula follows, which describes the evolution of such chains, *viz*, the Chapman-Kolmogorov equation(s)

$$\hat{P}^k = \hat{P}^{k-1} \cdot \hat{P} = \hat{P}^{k-l} \cdot \hat{P}^l; 0 < l < k \quad (42)$$

Obviously, Eq. (42) is a discrete analogue to Eq.(23) in describing the superposition of secondary 'wavelets'; and this holds independently of the fact, that one-step Markov chains with real-valued state variables describe diffusion-like processes (overdamped waves).

In discrete spaces, the principle of action-at-proximity means, that, during one time interval, only the next-neighbouring cells can be reached. A most important example of this class of Markov chains is constituted by the random walks (RW) (Spitzer 1964).

Consider the symmetric simple RW in 1D. An imaginary particle in an infinite chain of cells is supposed to hop at each step to one of the two neighboring cells, where the probability of hopping forward and backward equals one-half. The probability, $p_{k,i}$, to find the particle at time step k in cell i is given by the recursion formula.

$$p_{k,i} = \frac{1}{2} (p_{k-1,i-1} + p_{k-1,i+1}), \quad (43)$$

$$k = 0, 1, 2 \dots; -\infty < i < +\infty.$$

This is the well-known Euler forward scheme for the diffusion equation, $\partial T / \partial t = \partial^2 T / \partial x^2$.

The *f*undamental solution to Eq. (43) reads (the particle starts at $k = 0$ in cell $i = 0$).

$$p_{k,i}^f = \left\{ \begin{array}{l} 2^{-k} \binom{k}{n}; n \equiv \frac{k-|i|}{2} \text{ integer} \\ 0 \text{ half-integer} \end{array} \right\}. \quad (44)$$

This 'discrete Gaussian' is the analogue to the fundamental solution of the diffusion equation. The corresponding GF ("Green probability" in Keilson (1965), p. 80),

$$G_{k,i;k'i'} = p_{k-k',i-i'}^f, \quad (45)$$

possesses the Markov property (42).

$$G_{k,i;k'i'} = \sum_{i''} G_{k,i;k''i''} G_{k'',i'';k'i'}; k \geq k'' \geq k'. \quad (46)$$

In other words, the conditional or transition probabilities, which connect the different states of an one-step Markov chain, are at once the matrix elements of the GF of the corresponding difference equation of first order in the time parameter.

Pascal's triangle is a simple, but instructive example of Markovian 'number diffusion' obeying Huygens' recipe of construction (de Cogan & Enders 1991).

Accounting – analogously to the continuous case – for the actual boundary conditions, one may term discrete Huygens propagators such GF, or – more generally – evolution operators, which, per one time step, connects solely next-neighboring states and posses the Markov property (46).

B. A two-step Markov chain (random walk with correlation)

For first-order processes, such as the simple random walk, the Huygens propagator proves to be identical with the GF of the difference equation. This perfectly parallels the continuum case, where the Chapman-Kolmogorov equation (23) is fulfilled by the GF of equations of first-order in the time variable. Correspondingly, the GF of a multi-step equation of motion is, in general, not a (discrete) Huygens propagator. To get such one, one has 'to return' to a system of one-step equations of motion for a complete set of independent dynamical variables.

As an example, consider the partial difference equation of 2nd order

$$V_{k+2,i} = (V_{k+1,i-1} + V_{k+1,i+1}) + (\rho^2 - \tau^2)V_{k,i}, \quad (47)$$

being a discrete analogue to the telegraph equation; the parameters τ and ρ are determined by material parameters and mesh sizes. As such it has been proposed by Goldstein (1951) to model diffusion without the artefact of infinite speed of propagation. Du Fort & Frankel (1953) have shown that it realizes an explicit, but *unconditionally stable* finite difference routine for numerically solving the 1D diffusion equation. For $\rho = 0$, one gets Lax' scheme (1954) for hyperbolic equations of first order, cf eqs.(49) below. It also describes travelling voltage pulses on a network of lossless transmission lines and resistors, and the passivity of this network explains the stability of this scheme (Johns 1977, de Cogan 1998). The corresponding GF (Enders & de Cogan 1993) is not a Huygens propagator, because it does not obey the Chapman-Kolmogorov equation (46).

In the case of spatially variable material parameters, the determination of τ and ρ in Eq.(47) is not unique (Du Fort & Frankel 1953, Zauderer 1989). This determination can be made unique, when working with a system of first-order equations. Thus, following D'ALEMBERT, the field V may be decomposed into a left-running part, R , and a right-running part, L .

$$V_{k,i} = R_{k,i} + L_{k,i}. \quad (48)$$

Within a probabilistic treatment, $R_{k,i}(L_{k,i})$ is the probability to arrive from the right (left) at node i at step k . Within a network approach, the probabilities are replaced with traveling voltage pulses. This leads to the following system of two coupled partial difference equations of first order (Goldstein 1951, Johns 1977, Zauderer 1989),

$$L_{k+1,i} = \tau^L L_{k,i-1} + \rho^R R_{k,i-1}, \quad (49)$$

$$R_{k+1,i} = \rho^L L_{k,i+1} + \tau^R R_{k,i+1}. \quad (50)$$

Here, in dependence of the choice of how $\rho^{R,L}$ and $\tau^{R,L}$ vary from cell to cell, one obtains in the continuum limit the forward or backward Kolmogorov equation (Zauderer 1989) or Fick's second law (Johns 1977, Enders & de Cogan 1992).

In matrix form, eqs. (49) read $(\Delta_{\pm} R_{k,i\pm 1})$

$$\begin{pmatrix} \vec{L}_{k+1} \\ \vec{R}_{k+1} \end{pmatrix} = \begin{pmatrix} \tau^L \Delta_- & \rho^R \Delta_- \\ \rho^L \Delta_+ & \tau^R \Delta_+ \end{pmatrix} \begin{pmatrix} \vec{L}_k \\ \vec{R}_k \end{pmatrix} \equiv \hat{D} \begin{pmatrix} \vec{L}_k \\ \vec{R}_k \end{pmatrix}. \quad (51)$$

This is the two-step analogue to Eq. (39). $G_{k,i;k',i'} = (\hat{D}^{k-k'})_{i,i'}$ is the GF of eqs. (49).

C. Proper Huygens propagators

For $\rho^L = \rho^R = \rho = const$ and $\tau^L = \tau^R = \tau = const$, the system (51) is formally equivalent to Eq. (47). In this case – in agreement with the Cayley-Hamilton theorem – the eigenvalue equation of the transition matrix \hat{D} reads

$$\hat{D}^2 = \tau(\Delta_- + \Delta_+)\hat{D} + (\rho^2 - \tau^2)\hat{1}. \quad (52)$$

This corresponds to the diagonalization of the system (51) into the form of Eq. (47) for both \vec{L} y \vec{R} .

It is proposed to term the corresponding Huygens propagators, such as $(\hat{D}^{k-k'})_{i,i'}$, *proper* or *irreducible*, since the elements of them obey the multi-step equation of motion, too. This is an important property, because in this case, the eigenvalue equation of the transition matrix \hat{D} , diagonalizes the first-order equations of motion to a physically relevant equation.

A counter-example are the difference equations relating (\vec{V}_{k+1}, V_k) to (\vec{V}_{k-1}, V_{k-2}) .

Conjeture 2 *The eigenspectrum of the proper Huygens propagator approximates the eigenspectrum of the object under investigation.*

This would foster the fact, that the discrete formulation of HP yields construction principles for numerical algorithms for a wide variety of problems (cf Hoefer 1991). In particular, it would largely simplify the computation of eigenmodes by means of the Cayley-Hamilton theorem.

Conjeture 3 *In d dimensions, the discrete modelling of scalar wave propagation is related to a 2d-step Markov*

chain. In turn, in a n -step Markov chain, each step corresponds to 1 degree of freedom of motion, and vice versa: A motion with n degrees of freedom (n independent dynamical variables) can be mapped onto a n -step Markov chain.

The validity of this hypothesis would have important consequences for the understanding of complex systems (Ebeling 1992).

IX. FRACTIONAL FOURIER TRANSFORMATION AND HUYGENS' PRINCIPLE

The fractional Fourier transformation (FracFT) has been introduced as generalization of the all-present ordinary Fourier transformation (FT) with interesting applications in quantum mechanics (Namias 1980), optics (Lohmann 1993), and other fields. Its advantages for the description of optical propagation has been stressed by Alieva *et al* (1994). One may wonder, why its relationship to HP as a basic principle of optical propagation has not yet established. This connection multiplies the advantages noted by Alieva *et al* (1994) and adds new ones.

The FracFT of a function $f(x)$ is the function

$$F_\alpha(u) = \mathcal{F}^\alpha f(x) = \int_{-\infty}^{+\infty} f(x)K_\alpha(x,u)dx, \quad (53)$$

with the kernel ($n = 1; 2; \dots$)

$$K_\alpha = \begin{cases} \sqrt{\frac{1-i \cot(\alpha)}{2\pi}} e^{i(x^2+u^2) \cot \frac{\alpha}{2} - i x u \csc(\alpha)}; & \alpha \neq n\pi \\ \delta(x-u); & \alpha = 2n\pi \\ \delta(x+u); & \alpha = (2n-1)\pi \end{cases} \quad (54)$$

This kernel is continuous in the generalized function sense, $\mathcal{F}^{2n\pi}$ is the identity, and $\mathcal{F}^{2n\pi+\pi/2}$ is the ordinary FT. Almeida (1993, 1994) has derived the group property $\mathcal{F}^\alpha \mathcal{F}^\beta = \mathcal{F}^{\alpha+\beta}, i.e,$

$$K_{\alpha+\beta}(x,z) = \int_{-\infty}^{+\infty} K_\alpha(x,u) K_\beta(u,z)du. \quad (55)$$

The isomorphism with the Chapman-Kolmogorov equation (23) is obvious. In fact, up to a phase factor, $K_\alpha(x,u)$ is equivalent to the GF of the time-dependent Schroedinger equation for the harmonic oscillator (Agarwal & Simon 1994).

Generally speaking, the kernel of any transformation satisfying a relation like (55) is equivalent to the GF of a parabolic differential equation (such as the paraxial wave equation) in appropriate coordinates and, consequently, describes the propagation of some field. Therefore, the question arises, whether there are useful generalizations of the FracFT through choosing for the kernel other propagators, than that for the harmonic oscillator or for the parabolic index profile. In other words, are there further potential functions or index profiles which yield integral transformations with similarly useful properties as those of the Fourier and fractional Fourier transformations? Furthermore, are there applications for the generalization

of Eq. (55) to matrix functions as kernels of integral transformations?

From a computational point of view, it may be favourable to have got a discrete formulation of this theory. According to the foregoing section, this should be possible in terms of Markov chains or (transmission-line) networks. This could open a novel approach to wave-optical computations.

X. GENERAL THESIS FOR DESCRIBING PROPAGATION PROCESSES

The following thesis are proposed to built a starting point for an extension of the definition and application of HP to all propagation phenomena, which can be described through linear explicit differential and difference equations, respectively.

1. Propagation via action-at-proximity proceeds such, that the field excites secondary sources, which re-irradiate the field accordingly to the actual boundary and continuity conditions. Topologically, this principle applies on structures with next-neighbour interaction (local theories; cellular automata; certain coupled maps).
2. The propagating field is represented by a set of f independent (but interacting, of course) dynamical variables, where f equals the number of time-derivatives in the governing equations. In general, there are several such sets. A complete set obeys a system of f differential and difference equations of first order, respectively. Examples are the right-and left-running waves in Eq.(48), the wave amplitude and its "inner" speed of change, or field and flux density [eqs. (24) to (27)]. The flux density may play the role of the time-derivative of the field as independent dynamical variable, while its vector components are not independent of each other dynamical variables. A set of f one-step Markov chains provides the appropriate form for a discrete model of the propagation of f independent variables.
3. The (matrix-valued) GF of such a system contains the propagation of that complete set. It represents HP in the sense of action-at-proximity and superposition of secondary wavelets by means of the Chapman-Kolmogorov equation. In order to avoid perturbing boundary terms and to completely represent the propagation problem under consideration, the GF should fulfill the boundary conditions for the field variables in appropriate form. For such GF the term Huygens propagator is proposed.
4. The elimination of backward motion and the conservation of sharp, non-spreading fronts during propagation are special cases, that emerge naturally from the governing equations and do not need additional assumptions.

XI. CONCLUSIONS

For Feynman (1948), HP was – in geometrical-optical formulation – valid for matter waves, since the

Schroedinger equation is of first order in time; the Chapman-Kolmogorov equation holds true for the amplitudes of the quantum-mechanical transition probabilities. Schroedinger (1926b) has extensively quoted HP "in its old, naive form, not in the rigorous Kirchhoff form", ie, in the same meaning that was understood by Feynman. As its expression he has seen the Hamilton-Jacobi equation – an equation of first order in time. The representation of HP proposed in this paper unifies the description of propagation processes modelled by parabolic and hyperbolic differential equations. It is the same one for geometrical and for wave optics; the former being a limit case, but without ad-hoc assumptions.

The mathematical formulation of HP in form of the Chapman-Kolmogorov equation (23) implies the following important conclusions.

- (i) Huygens' construction can be applied to spreading wave fronts as well.
- (ii) Wave propagation is a Markov process (speech recognition bases on this fact).
- (iii) HP in that sense holds true for Dirac and similar quantum fields as well as for diffusion processes.

The thesis of section 10 are deliberately formulated in such a general manner, that they apply, among others, also to the cases of spatial anisotropy (birefringence, Huygens 1690), isotropy in the sense of the (local) line element (Schroedinger 1926a,b), nonlinear and fluctuating propagation conditions (Vanneste et al. 1992, Enders 1993), audio-holography (Illenyi & Jessel 1991), and the states in electrical power systems (Vasin 1990). HP needs no correction as proposed by Miller (1991), and the difficulties discussed by Johns (1974) are lifted as well. The mathematical representation of HP by means of propagators and the Chapman-Kolmogorov equation throws also new light upon the relation between the fractional Fourier transformation and wave propagation and suggests further generalizations and applications in this field.

Schwartz (1987) wrote, Physically this [HP] makes no sense at all. Light does not emit light; only accelerating charges emit light." Indeed, not the wavefront itself irradiates the secondary wavelets, but the matter it excites (including the so-called vacuum exhibiting finite values ϵ_0 and μ_0) does so. Since HP is not concerned with the mechanisms of excitation and re-irradiation, the GF and, thus, the Chapman-Kolmogorov equation cope with this physical point.

The following text resembles Feynman's (1948) original statement and fosters the view on HP presented in this contribution. "Huygens principle follows formally from the fundamental postulate of quantum electrodynamics – that wavefunctions of every object propagate over any and all allowed (unobstructed) paths from the source to the given point. It is then the result of interference (addition) of all path integrals that defines the amplitude and phase of the wavefunction of the object at this given point, and thus defines the probability of finding the object (say, a photon) at this point. Not only light quanta (photons), but electrons, neutrons, protons, atoms, molecules, and all other objects obey this simple

principle." ([http://en.wikipedia.org/wiki/Huygens' principle](http://en.wikipedia.org/wiki/Huygens%27_principle), Feb. 12, 2008).

Difference equations representing a discrete HP are directly suited for computing all propagation processes that can be modelled through explicit differential equations. This should enable the simultaneous and self-consistent computation of interacting fields of different type, eg, heat diffusion and electromagnetic waves in lasers (Enders 1992), in microwave ovens or in lenses and mirrors for high-power beams. Within explicit schemes, self-consistency can be achieved at every (time) step, whereby convergency is considerably accelerated.

One of such numerical algorithms is the Transmission-line Matrix Modeling Method (TLM), an explicit finite-difference scheme describing travelling voltage pulses on a mesh of lossless transmission lines and lumped resistors (Christopoulos 1995, de Cogan 1998, de Cogan et al 2005). These difference equations trace a practically realizable physical process obeying HP, too. Due to that, a TLM routine exhibits excellent stability properties, which – among others – are exploited in commercial programm packages. The GF of the coupled one-step TLM equations is a proper Huygens propagator exhibiting the computational advantages described above (cf Enders & Włodarczyk 1993). Johns' (1987) symmetric condensed node for solving Maxwell's equations in 3D obeys even Hadamard's Minor Premise (Johns & Enders 1992).

Delsanto and coworkers (1992) have stressed that a local interaction approach to simulation is favored by three practical advantages:

- (i) extremal speed due to immediate parallelizability;
- (ii) complex problems can be treated in a simple manner, since the local internodal connections are arbitrarily variable;
- (iii) the same code can be used for quite different problems, since the iterations (difference equations) are principally, ie, up the the values of the coefficients, identical. Such algorithms belong to the class of cellular automata (Wolfram 1986), where there is no limitation for the state set of the nodes.

Thus, when "the purpose of computing is insight, not numbers" (Hamming 1973), then an approach basing on a discrete HP is an ideal starting point for the development of codes, not at least due to its philosophy of modeling (Johns 1979, Toffioli 1984, Vichniac processes by means of clear division into elementary steps, which in turn display a large variety of behavior, may contribute to the unity of the treatment of propagation phenomena in different environmental conditions.

ACKNOWLEDGMENTS

These investigations revise and largely extend the earlier review (Enders 1996), the work on which had been started in 1991 at the University of East Anglia, Norwich, UK, with Dr. de Cogan and which had been inspired from various discussions with Prof. W. Brunner, Prof. W. Ebeling (1992), Prof. T. Elsaesser, Dr. H. Engel, Dr. R. Guether, D. P. Johns (Johns & Enders 1992), Msc, Dr. R.

Mueller, Prof. H. Paul, Dr. A. J. Wlodarczyk (Enders & Wlodarczyk 1993) and Dr. M. Woerner. Then I have learned from Dr. C. Vanneste how to account correctly for the incoming waves, so that HP for TLM could be established in general form, both locally (Enders & Vanneste 2003) and globally (Enders 2001). Section 2 emerged from my collaboration with Dr. D. Suisky on Euler's representation of classical mechanics and its axiomatic generalization to quantum mechanics (Enders 2006). Most about the Helmholtz decomposition I have learned from papers by Prof. O. Keller (for a review, see Keller 2005) and from various personal explanation by him. Last but not least I'm indebted to Dr. C. Francis and other posters in the moderated Usenet group sci.physicsFOUNDATIONS for their clarifying remarks. This work was partially supported by the Deutsche Akademie der Naturforscher Leopoldina (Enders 1996a).

REFERENCES

- [1] Agarwal, G.S. & Simon, R. 1994 A simple realization of fractional Fourier transform and relation to harmonic oscillator Green's function. *Opt. Commun.* 110, 23-26.
- [2] Almeida, L.B. 1993 Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing, Minneapolis, Minnesota, April 1993 (New York: IEEE).
- [3] Almeida, L.B. 1994 An introduction to the angular Fourier transform. *IEEE Trans. Signal Process.* 33, 6182-6187.
- [4] Alieva, T., Lopez, V., Agullo-Lopez, F. & Almeida, L.B. 1994 The fractional Fourier transform in optical propagation problems. *J. Mod. Optics* 41, 1037-1044.
- [5] Barr, A.H. 1991 *Physically-Based Modeling in Computer Graphics and Animation*. Addison-Wesley, Reading (Mass.).
- [6] Chapman, S. & T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases. An account of the kinetic theory of viscosity, thermal conduction, and diffusion in gases*, Cambridge: Cambridge Univ. Press 1939.
- [7] Chopard, B. & Droz, M. 1998 *Cellular Automata Modeling of Physical Systems*. Cambridge, Cambridge Univ. Press.
- [8] Christopoulos, C. 1995 *The Transmission-Line Matrix Method – TLM*. Oxford, Oxford Univ. Press.
- [9] Courant, R., Friedrichs, K. & Lewy, H. 1928 *Über die partiellen Differenzgleichungen der Mathematischen Physik*. *Math. Annalen* 100, 32-74; Engl. transl.: *IBM J. Res. Develop.* 11 (1967) 215-234.
- [10] Courant, R. & Hilbert, D. 1968 *Methoden der Mathematischen Physik II*. Springer, Berlin/Heidelberg/New York, 2nd ed.
- [11] De Cogan, D. 1998 *Transmission Line Matrix (TLM) Techniques for Diffusion Applications*. Gordon & Breach.
- [12] De Cogan, D., O'Connor, W.J. & Pulko, S. 2005 *Transmission Line Matrix in Computational Mechanics*. Routledge, New York.
- [13] De Cogan, D. & Enders, P. 1991 Microscopic effects in TLM heat flow modelling. *IEE Digest No. 1991/157*, pp. 8/1-8/11.
- [14] DeSanto, J.A. 1992 *Scalar Wave Theory. Green's Functions and Applications*. Springer, New York etc.
- [15] Delsanto, P.P., Chaskelis, H.H., Mignogna, R.B., Whitcombe, T.V. & Schechter, R.S. 1992 *Connection Machine Simulation of Ultrasonic Wave Propagation: Two Dimensional Case*. *Rev. Progr. Quant. Nondestr. Eval.* 11 (D.O. Thompson & D.E. Chimenti, Eds.), Plenum Press, New York, pp.113-120.
- [16] Dirac, P.A.M. 1949 *Forms of Relativistic Dynamics*, *Rev. Mod. Phys.* 21, 392-399.
- [17] Du Fort, E.C. & Frankel, S.P. 1953 *Stability Conditions in the Numerical Treatment of Parabolic Differential Equations*. *Math. Tables and other Aids to Computing* 7, 135-152.
- [18] Dynkin, E.B. 1965 *Markov Processes*. Springer, Berlin/Göttingen/Heidelberg (2 Vols.).
- [19] Dyson, F. J. 1993 *George Green and physics*. *Physics World*, Aug., 33-38.
- [20] Ebeling, W. 1992, Talk at the Landsberg-Symposium, Berlin, Aug. 8 (unpubl.), and private communication.
- [21] Einstein, A. 1917 *Eine Ableitung des Theorems von Jacobi*. *Sitzber. Kgl. Preuss. Akad. d. Wiss., phys.-math. Kl.* XLVI, 606-608.
- [22] Enders, P. 1992 *The Chapman-Kolmogorov Equation as Representation of Huygens' Principle and the Monolithic Self-Consistent Numerical Modelling of Lasers*. Proc. 4th Int. Conf. Physics Computing '92, Prague, Aug. 24-28, R.A. de Groot & J. Nadrchal (Eds.), World Scientific, Singapore etc. (pp. 328-329).
- [23] Enders, P. 1993 *Comment on "A wave automaton for time-dependent wave propagation in random media" by C. Vanneste et al.* *europ. phys. lett.* 21, 791-793.
- [24] Enders, P. 1996a *A new view on Huygens principle*, OSA Annual Meeting, Rochester 1996, Paper WGG50.
- [25] Enders, P. 1996b *Huygens' Principle and the Modelling of Propagation*. *Eur. J. Phys.* 17 (1996) 226-235.
- [26] Enders, P. 2001 *Huygens' principle in the transmission line matrix method (TLM). Global theory*. *Int. J. Num. Modell.: Electronic Networks, Dev. and Fields* 14, 451-456.
- [27] Enders, P. 2006 *Von der klassischen Physik zur Quantenphysik. Eine historisch-kritische deduktive Ableitung mit Anwendungsbeispielen aus der Festkörperphysik*. Berlin · Heidelberg: Springer.
- [28] Enders, P. & de Cogan, D. 1992 *Correlated random walk and discrete modelling of propagation through inhomogeneous media*. Proc. 4th Int. Conf. Physics Computing '92, Prague, Aug. 24-28, R.A. de Groot & J. Nadrchal (Eds.), World Scientific, Singapore etc. (pp. 307-308).
- [29] Enders, P. & de Cogan, D. 1993 *The efficiency of transmission-line matrix modeling – a rigorous viewpoint*. *Int. J. Num. Modell.: Electron. Networks, Dev. and Fields* 6, 109-126.
- [30] Enders, P. & Vanneste, C. 2003 *Huygens' principle in the transmission line matrix method (TLM). Local theory*. *Int. J. Num. Modell.: Electronic Networks, Dev. and Fields* 16, 175-178.

- [31] Enders, P. & Wlodarczyk, A. J. 1993 Towards Exactly Modelling Open/Absorbing Boundaries. 2nd Int. Workshop Discrete Time Domain Modelling of Electromagnetic Fields and Networks, Berlin, 28.-29.10.93 (unpublished).
- [32] Feynman, R.P. 1948 Space-Time Approach to Non-Relativistic Quantum Mechanics. Rev. Mod. Phys. 20, 367-387.
- [33] Feynman, R.P. & Hibbs 1965 Quantum Mechanics and Path Integrals. New York: McGraw-Hill.
- [34] Goldstein, S. 1951 On diffusion by discontinuous movements, and on the telegraph equation, Quart. Journ. Mech. and Appl. Math. IV, 129-156.
- [35] Guenther, P. 1988 Huygens' Principle and Hyperbolic Equations. Academic Press, New York.
- [36] Hadamard, J. 1953 The Cauchy Problem and the Linear Hyperbolic Partial Differential Equations. Dover, New York 1953.
- [37] Hamming, R.W. 1973 Numerical methods for scientists and engineers. McGraw-Hill, New York, 2nd ed.
- [38] H. Helmholtz, Uber Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen, J. Reine Angew. Math.55 (1858) 25-55.
- [39] Hoefler, W.J.R. 1991 Huygens and the Computer – A Powerful Allianz in Numerical Electromagnetics. Proc. IEEE 79, 1459-1471.
- [40] Hoefler W.J.R & So P.P.M. 1991 The Electromagnetic Wave Simulator. A Dynamical Visula Electromagnetics Laboratory based on the Two-Dimanesional TLM Method. Wiley: Chichester etc.
- [41] Huygens, CHR. 1669 De motu corporum ex percussione, London: Royal society, Talk Jan. 4, 1669, publ. 1703.
- [42] Huygens, Chr. 1673 Horologium oscillatorium sive de motu pendulorum ad horologia aptato demonstrationes geometricae, Paris 1673; in: The complete Works of Christiaan Huygens, Vol. XVII, Den Haag 1934.
- [43] Huygens, CHR. 1690 Trait'e de la lumiere. Pierre van der Aa, Leiden.
- [44] Illenyi, A.& Jessel, M. 1991 Holographics, a spread-out of the basic ideas on holography into audio-acoustics. In: Holography, Commemorating the 90th Anniversary of the Birth of Dennis Gabor, Tatabanya, H, 2-5 June 1990; SPIE Institutes for Advanced Optical Technologies, Vol. IS 8 (1991), Bellingham: SPIE Optical Engineering Press, pp. 39-52.
- [45] Iwanenko, D.I. & Sokolow, A. 1953 Klassische Feldtheorie. Akademie-Verlag, Berlin (p. 78)
- [46] Johns, D.P. & Enders, P. 1992 Wave Optics Computing by a Network-Based Vector Wave Automaton. Proc. 4th Int. Conf. Physics Computing '92, Prague, Aug. 24-28, R.A. de Groot & J. Nadrchal (Eds.), World Scientific, Singapore etc. (pp. 359-361).
- [47] Johns, P.B. 1974 A new mathematical model to describe the physics of propagation. The Radio and Electronic Engin. 44, 657-666.
- [48] Johns, P.B. 1977 A simple explicit and unconditionally stable numerical routine for the solution of the diffusion equation. Int. J. Numer. Meth. Engin. 11, 1307-1328.
- [49] Johns, P.B. 1979 The art of modelling. Electronics and Power 25, 565-569.
- [50] Johns, P.B. 1987 A Symmetrical Condensed Node for the TLM Method. IEEE Trans. Microwave Theory Tech. MTT-35, 370-377.
- [51] Keilson, J. 1965 Green's Function Methods in Probability Theory. Griffin & Co., London.
- [52] Keller, O. 2005 On the theory of localization of photons, Phys. Rep. 411, 1-232.
- [53] Kolmogorov, A.N. 1931 Uber die analytischen Methoden der Wahrscheinlichkeitsrechnung, Math. Ann. 104, 415ff.
- [54] Kolmogorov, A.N. 1933 Grundbegriffe der Wahrscheinlichkeitsrechnung, Berlin: Springer.
- [55] Lax, P.D. 1954 Week solutions of Nonlinear Hyperbolic Equations and Their Numerical Computation. Comm. Pure Appl. Math. 7, 159-193.
- [56] Lohmann, A. 1993 Image rotation, Wigner rotation, and the fractional Fourier transform. JOSA A 10, 2181-2186.
- [57] Longhurst, R.S. 1973 Geometrical and Physical Optics. Longman, London and New York, 3rd ed.
- [58] Lukasiewicz, J. 1957 Aristotle's Syllogistic From the Standpoint of Modern Formal Logic. Oxford: The Clarendon Press.
- [59] Macke, W. 1958 Wellen. Geest & Portig, Leipzig, Section 443.
- [60] Mie, G. 1941 Lehrbuch der Elektrizit'at und des Magnetismus. Eine Experimentalphysik des Welt'athers f'ur Physiker, Chemiker, Elektrotechniker, Stuttgart: Enke (2nd ed.).
- [61] Miller, D.A.B. 1991 Huygens's wave propagation principle corrected. Optics Lett. 16, 1370-1372.
- [62] Mueller, I. 1967 Zum Paradoxen der W'armeleitungstheorie. Z. Physik 198, 329-344.
- [63] Naas, J. & Schmid, H.L. 1974 Mathematisches W'orterbuch. Akademie-Verlag, Berlin, and Teubner, Leipzig, 3rd ed., Vol. I.
- [64] Namias, V. 1980 The fractional Fourier transform and its application in quantum mechanics. J. Inst. Math. Appl. 39, 241-265.
- [65] Scharf, G. 1994 From Electrostatics to Optics. Springer, Berlin etc.
- [66] Schroedinger, E. 1926a Quantisierung als Eigenwertproblem. Erste Mitteilung. Ann. Physik [4] 79, 361-376; 1926b Zweite Mitteilung. Ibid. 489-527; 1926c Dritte Mitteilung: St'orungstheorie, mit Anwendung auf den Starkeffekt der Balmerlinien. Ibid. 80, 437-490; 1926d Vierte Mitteilung. Ibid. 81, 109-13.
- [67] Schwartz, M 1987 Principles of Electrodynamics. Dover, New York (quoted after <http://www.mathpages.com/home/kmath242/kmath242.htm>).
- [68] Simonyi, K Kulturgeschichte der Physik, Leipzig u.s.w.: Urania 1990.
- [69] Sommerfeld, A. 2001 *Vorlesungen über theoretische Physik*, Bd. III *Elektrodynamik*, Frankfurt a. Main: Deutsch (4th ed.).
- [70] Spitzer, F. 1964 Principles of Random Walk. Van Nostrand, Princeton (N.J.).

- [71] Toffioli, T. 1984 Cellular Automata as an Alternative to (Rather Than an Approximation of) Differential Equations in Modeling Physics. *Physica* 10D, 117-127.
- [72] Vanneste, C., Sebbah, P. & Sornette, D. 1992 A Wave Automaton for Time-Dependent Wave Propagation in Random Media. *Europhys. Lett.* 17, 715-720.
- [73] Vasin, V. P. 1990 Methods of global analysis of electric power systems' states of operation. *Power Engineering (New York)* 28, 22-34
- [74] Vichniac, G.Y. 1984 Simulating Physics with Cellular Automata. *Physica* 10D, 96-116.
- [75] West, B.C. 1992 The use of fractals in optical wave propagation through random media. *Proc. Conf. Lasers and Electro-Optic CLEO'92*, 10.-15.05.1992, Anaheim, CA, S. 506.
- [76] Wiener, J. 1995 ARISTOTLE'S SYLLOGISM: LOGIC TAKES FORM. <http://www.perseus.tufts.edu/GreekScience/Students/Jordana/LOGIC.html>.
- [77] Wigner, E.P. 1963 Events, laws of nature, and invariance principles. In: G. Ekspong (Ed.) 1998 Nobel Lectures in Physics Vol 4 1963-1980, Singapore: World Scientific, p.6-17.
- [78] Wolfram, S. 1986 Theory and Applications of Cellular Automata. World Scientific, Singapur.
- [79] Zauderer, E. 1989 Partial Differential Equations of Applied Mathematics, Wiley, New York.