Studies in the Generalized Theory of Gravitation II:
The Velocity of Light

BY

O. HITTMAIR

AND

E. SCHRÖDINGER

STUDIES IN THE GENERALIZED THEORY OF GRAVITATION II: THE VELOCITY OF LIGHT

By

O. HITTMAIR and E. SCHRÖDINGER.

1. The various meanings of the velocity of light.

As is well known, the influence of a gravitational field on the propagation of light was in Einstein's theory of gravitation one of the main issues, which by good luck was just within the reach of careful observation, that decided in favour of this theory. In its recent non-symmetric generalization the question arises, how the skew field, which is tentatively regarded as the electromagnetic field, influences the propagation of waves of this very field itself, that is of light, if our tentative view is appropriate. The subject is touched upon in several recent papers[1]. Pondering these remarks we found that the first requirement is a revision of

the concept "velocity of light" which has two fundamental meanings in the older theory, but three in the new one. We shall discuss both cases briefly, beginning, of course, with the older one.

---

[1]
R. L. Ingraham, Annals of Math., 52, 743, 1950.

P. Udeschini, Rend. Lincei, 9, 256, 1950; 10, 21, 121, 390, 1951.

To declare that the symmetric tensor-field g_{ik} is the world metric, implies the physical assumptions that we have unit measuring rods and clocks which measure the invariant interval

$$ds^2 = g_{ik}\, dx_i\, dx_k$$

in particular cases, viz. a rod for simultaneous positions of its end-points (irrespective of its orientation), and a clock, when at rest, for the two world points indicated by its having advanced by one unit. - A second and independent physical hypothesis is about the propagation of light: the dx_k for two neighbouring points reached by <u>one</u> light signal shall satisfy $ds^2 = 0.$

If we lay out a local frame with the help of such rods and clocks, choosing rectangular Cartesian space coordinates, the g_{ik} acquire the values $g_{11} = g_{22} = g_{33} = -1$, $g_{44} = 1$, all others zero (we shall call this "Galilean"), and the velocity of light becomes 1 in all directions; all this in virtue of our assumptions, which would have to be discarded if experience contradicted. Needless to say that in this local frame there can be no question of rods or clocks changing, contracting or being retarded, etc., by the gravitational field or by orient-ation, etc.

In an extended gravitational field one can always adopt a general world frame of which this local frame at a given world-point forms part, and usually, with sufficient accuracy, in a wide neighbourhood (it <u>may</u> be a geodesic frame, but that is not the point; the one we use every day is not).

This is the <u>first</u> meaning of the velocity of light in the

older theory.   It is the only invariant meaning, although -

nay, because - it refers to a very special frame.   However

the very nature of a true gravitational field makes it im-

possible to choose a world-frame that produces this simple

state of affairs everywhere.   We cannot avoid using a gen-

eral frame for studying extended phenomena such as the deflec-

tion of light passing near the limb of the sun, or the wave-

length of a spectral ray, emitted on a white dwarf and meas-

ured in a terrestial observatory.   One can as a rule - and

does, of course, if one can - avoid $g_{i4} \neq 0$ (i = 1, 2, 3).

This leads to

$$ds^2 = g_{44}dx_4{}^2 - \sum_1^3 \sum_1^3 g_{ik}dx_i dx_k \quad .$$

This induces one to interpret a $g_{44} \neq 1$ as a change of rate

of the clock, and non-Galilean spatial g_{ik}'s as a change of

length of the rod, depending also on its orientation, and

finally to say that the velocity of light is $\neq 1$ and is

"anisotropic".   But it is quite clear that all these notions

are eminently non-invariant and locally meaningless, since

they disappear in the local frame.   In fact you may, if you

are provoked, produce all kinds of freak by choosing a suit-

ably unsuitable frame !

    This, then, is the second, the non-invariant concept of

the velocity of light.   That it is sometimes needed for pre-

dicting very definite and substantial phenomena, is well

known.

    In the way of a digression I should like to repudiate a

third meaning, used by Udeschini l. c. and hailing, it seems,

from Levi-Cività, viz. to regard $\sqrt{g_{44}}$ as the velocity of light. The idea is, to regard in the above equation the double sum as the square of the distance, say dr^2, and dx_4 as the time interval so that $ds^2 = 0$ entails $\dfrac{dr}{dx_4} = \pm \sqrt{g_{44}}$. The physicist has no use for this concept. It combines the defects of being non-invariant and yet producing wrong results for the gravitational deviation of a light ray. -

We now turn to the generalized theory in which g_{ik} is non-symmetric. To regard its even part $\underline{g_{ik}}$ as the world-metric amounts to hold on to the assumptions about clocks and measuring rods, now with regard to the invariant

$$ds^2 = \underline{g_{ik}}\, dx_i\, dx_k.$$

In the local frame the $\underline{g_{ik}}$ will then turn out Galilean, as the g_{ik} did before. But what about the propagation of light? It consists of waves of $\overset{\vee}{g_{ik}}$, and the latter are controlled by the field equations just as much as the $\underline{g_{ik}}$. So there is now no room for an assumption as $ds^2 = 0$ or any other. If we wish to know how such waves are propagated we have to consult the field equations. These stipulate an intimate interaction between the $\underline{g_{ik}}$- and $\overset{\vee}{g_{ik}}$-fields. Hence, with a rapidly changing $\overset{\vee}{g_{ik}}$-field the $\underline{g_{ik}}$ could in general not remain Galilean for more than a split second!

It is, however, reasonable to define the "behaviour of light" in the following manner. We split the total skew $\overset{\vee}{g_{ik}}$ field additively into two parts: an infinitely weak, rapidly oscillating part that represents the light-wave whose behaviour we wish to investigate, and the remaining background-field, which we do not restrict as to its magnitude but which we

consider to vary slowly in space and time, so that we may regard it as constant in the neighbourhood of the point in question just as we do with the Galilean g_{ik}, whose distortion by the infinitely weak rapidly changing light wave we neglect. In this way we shall find

i) with the g_{ik} Galilean and no background-g_{ik} the propagation of light is normal, $ds^2 = 0$, the velocity of light being constant and $= 1$.

ii) With the g_{ik} Galilean and a background-g_{ik}, the latter modifies the propagation of light, owing to the non-linearity of the equations controlling g_{ik}. This behaviour may still be called local and invariant inasmuch as it is only an interplay of local fields and undergoes surveyable changes on Lorentz transformation.

iii) In a general world-frame the background-g_{ik}, by acting as a source of the gravitational field g_{ik}, has also indirect, non-local, non-specific influence on the non-invariant behaviour of light in such a frame; very generally speaking this influence is of the same type as in the older theory and raises no new problems.

These are the three meanings of the "velocity of light" in the new theory. The feature of specific interest is, of course, (ii), comprising (i) as a special case. We deal with it in the next section.

2. The influence of the local background-field.

It is easy to show[1] that the density $\mathfrak{f}^{ik}$ and the tensor g_{ik} are, if you allow for the generalized metric g_{ik}, in the same relation to each other as the 2 six-vectors in Born's non-linear electrodynamics; $\mathfrak{f}^{ik}$ is to be identified with the (B,E)-tensor, the g_{ik} with the (H,D)-tensor. Then the relation $\mathfrak{f}^{ik}_{,k} = 0$ - the only Maxwellian set that the unified theory yields[2] - corresponds to curl E + $\dot{\text{B}}$ = 0, div B = 0, while the other set is obtained by defining $g_{[ik,l]}$ as the 4-current and putting it zero where there is to be none. This yields curl H - $\dot{\text{D}}$ = 0, div D = 0 (generalized for the metric g_{ik}). Hence the local velocity of light (the only one that has an invariant meaning), since it is obtained by introducing a local Galilean frame (-1, -1, -1, 1), is governed precisely by Born's theory, and we can make use of results worked out previously.

We shall use the letters B, E, D, H for a field of arbitrary strength (the background field), that we think of as

---

1)
  E. Schrödinger, Proc. Roy. Ir. Acad. 51 (A), p. 214, 1948. In equation (5,6a) the factor h is missing in the second term of the 2nd member.
2)
  Equation (4,8b) l.c. and the suggestion (5,7) become void in the more stringent form of the theory, which we are using here. - Also, in the present connection, being interested in local phenomena, we discard the cosmological term altogether.

locally homogeneous and static or, at any rate, as varying slow-
ly in space and time compared with the vibrations of the light-
fields whose propagation we are about to investigate (but for
which no notation will be required).   In the case of a purely
electric field E (and the corresponding D, see below) it has
been found[3], that the velocity of propagation u in the direction
of the wave normal of a <u>weak</u> plane wave crossing the background
field is, irrespective of polarization and frequency, given by

$$u^2 = A^2 \sin^2\omega + \cos^2\omega, \qquad (1)$$

where ω is the angle between the wave-normal and E.   Thus the
field E produces anisotropy, but no double refraction, and it
does not upset the central symmetry ($\omega \longrightarrow \pi - \omega$).   The
scalar A, which we regard as positive, is

$$A = \sqrt{1 - E^2} . \qquad (2)$$

Its reciprocal plays the part of dielectric constant <u>for the</u>
<u>background-field</u>

$$D = A^{-1} E . \qquad (3)$$

It is not difficult to generalize the argument N. O. 1. c. so
as to embrace the case of <u>parallel fields</u> E and H.   The only
change is that then

$$A = \sqrt{1 - E^2 - H^2} \qquad (2a)$$

while (3) is supplemented by[4]

---

3)
    E. Schrödinger, Non-linear Optics, Proc. Roy. Ir. Acad.
47(A), p. 101 f.   (This paper shall be quoted as N. O.)
4) N. O. p. 81, equations (2,6)

$$B = A^{-1} H. \tag{3a}$$

To obtain the general case when E and H are not necessarily parallel, the simplest way is to put (1), <u>without changing its content</u>, into Lorentz-invariant form. We first express A^2 by the invariants of the (B,E)-tensor. From (2a) and (3a)

$$A^2 = 1 - E^2 - A^2 B^2$$

$$A^2 = \frac{1 - E^2}{1 + B^2} = \frac{2 - E^2 + B^2}{1 + B^2} - 1 = 1 - \frac{E^2 + B^2}{1 + B^2} \tag{4}$$

$$\frac{1 - A^2}{1 + A^2} = \frac{E^2 + B^2}{2 + B^2 - E^2} = \frac{I}{1 + I_1} \quad,$$

where we have put

$$I_1 = \tfrac{1}{2}(B^2 - E^2) \qquad\qquad I_2 = (BE) = \pm |B|\,|E| \tag{5}$$

$$I = \sqrt[+]{I_1{}^2 + I_2{}^2} \; .$$

From the last equation (4)

$$A^2 = \frac{1 + I_1 - I}{1 + I_1 + I} \quad . \tag{6}$$

Now we divide (1) by the square of the wavelength[*] λ, in order to introduce the covariant wave vector k_1, k_2, k_3, k_4, the latter being u/λ, while k_2 shall be the component in the direction

---

[*] One could avoid specializing in harmonic waves; but by doing so no harm is done and language is simplified.

of the background field ($\omega = 0$). Thus, from (1):

$$k_4^2 = A^2 (k_1^2 + k_3^2) + k_2^2, \tag{7}$$

Inserting the value of A^2 from (6) we easily obtain

$$(1 + I_1)(k_4^2 - k_1^2 - k_2^2 - k_3^2) = - I (k_1^2 + k_3^2 + k_4^2) + I k_2^2 \tag{8}.$$

It remains to put the second member into invariant form.
Using transitorily the <u>contravariant</u> components k^1:

$$(1 + I_1)k^1 k_1 = I k^1 k_1 - I k^2 k_2 + I k^3 k_3 - I k^4 k_4. \tag{8a}$$

We now introduce, as a purely mathematical tool the <u>conventional</u>
Maxwellian energy-momentum-stress tensor of the tensor (B,E),
and call it $4\pi T^1_m$. It is easy to see that in our special frame,
with both B and E parallel to the direction labelled 2, T^1_m is
diagonal with components

$$- I, \qquad I, \qquad - I, \qquad I.$$

Hence the invariant form of (8a) reads

$$(1 + I_1) k^1 k_1 = - T^1_m k^m k_1, \tag{8b}$$

and must, of course, hold in every frame; which means <u>inter
alia</u> for <u>any</u> field (B,E). Given this field, we shall still
try to use the most convenient frame. We cannot without loss
of generality make T diagonal, because there will be a "Poynting-
vector" unless B and E are parallel. But we can reduce that
P. V. to one component (say in the 1-direction) and take for
"2" and "3" the other two axes of the (threedimensional)

"stress-tensor".   It is not difficult to see that T^1_m is then the following array

$$\begin{pmatrix} -w & 0 & 0 & +\sqrt{w^2 - I^2} \\ 0 & I & 0 & 0 \\ 0 & 0 & -I & 0 \\ -\sqrt{w^2 - I^2} & 0 & 0 & w \end{pmatrix} \qquad . \qquad (9)$$

The letter w has been chosen for the "energy density".   This we introduce in (8b), which at the same time we multiply by the square of the wave-length and thus replace the (covariant) k_1 by

$$- \cos\Theta, \qquad - \sin\Theta\cos\phi, \qquad - \sin\Theta\sin\phi, \qquad u.$$

So Θ is the angle between the wave normal and the "Poynting-vector", ϕ is the azimuth around the latter direction.   Taking good care of the signs we get in this way from (8b) and (9)

$$(1 + I_1)(u^2 - 1) = 2u\cos\Theta\sqrt{w^2 - I^2} - w\cos^2\Theta \ +$$
$$+ I\sin^2\Theta\cos 2\phi - w u^2$$

or

$$u^2 - 2u\cos\Theta\,\frac{\sqrt{w^2 - I^2}}{1 + I_1 + w} + \frac{w\cos^2\Theta - I\sin^2\Theta\cos 2\phi - I_1 - 1}{1 + I_1 + w} =$$
$$= 0. \qquad\qquad\qquad (10)$$

This quadratic equation gives u, the velocity of propagation of the wave plane in the direction of the wave normal indicated

by the angles Θ, ϕ.   The most striking feature is the term

linear in u, which is bound up with the non-vanishing "Poynting-

vector".   It means that the centre of symmetry is lost; u has

in general different values for opposite directions $(\phi, \Theta) \longrightarrow$

$(\phi + \pi, \pi - \Theta)$.   Apart from extreme cases (see below) the

two roots have opposite signs; the negative root, with reversed

sign, gives u for the opposite direction.   The easiest way of

procuring a synoptic view of the rather intricate state of aff-

fairs is to construct the eikonal - the envelope, after unit

time, of all the wave planes that have passed the origin simul-

taneously in all possible directions.   We skip the proof that

this eikonal __is__ an ellipsoid in standard orientation, but with

its centre displaced in the 1-direction.   Taking this for

granted we easily find the half-axes and the displacement by

computing u in the directions of the coordinates.   In this

way one obtains for the half-axes

$$
\underset{(1)}{\sqrt{\dfrac{1 + 2I_1 - I_2{}^2}{1 + I_1 + w}}}
\qquad
\underset{(2)}{\sqrt{\dfrac{1 + I_1 + I}{1 + I_1 + w}}}
\qquad
\underset{(3)}{\sqrt{\dfrac{1 + I_1 - I}{1 + I_1 + w}}}
\tag{11}
$$

and for the displacement of the centre in the direction of the

"Poynting-vector" (+1-direction)

$$
+ \; \frac{\sqrt{w^2 - I^2}}{1 + I_1 + w} \; .
\tag{12}
$$

In general the three half-axes are different.   Rotational

symmetry around the 2-axis occurs for $w = I$; this is the

special case of B and E parallel, from which we started;
u = 1 for the ± 2-directions.   The ellipsoid is "prolonged".
The displacement is zero.   – A flattened ellipsoid of rota-
tion around the 1-axis is obtained when both invariants
vanish: $I_1 = I_2 = I = 0$.   The displacement (12) sub-
sists in this case, and cannot be transformed away.   The
velocity u in the direction of the "Poynting-vector" is un-
changed

$$u = \frac{1}{1 + w} + \frac{w}{1 + w} = 1.$$

It is the case of E and B orthogonal and equal, and from (6)
also equal to D and H.   It is almost a gift from heaven,
due obviously to continuity, that this singular case is em-
braced, though it seems out of reach of the considerations
that led to equation (10).

Since a plane wave proceeding with the maximum velocity
u = 1 does so in every Lorentz frame, there must also in the
general case be just two directions in which (but not in the
opposites!) u = 1.   (For B ‖ E these are the directions
± 2; in the singular case they happen to coincide in the
direction + 1),   Obviously we have to seek them in the
(1,2)-plane, i. e. for $\phi = 0$ or Π .   So we have cos 2ϕ
= 1 in equation (10), which we write

$$(1 + I_1 + w) u^2 - 2u \cos\Theta \sqrt{w^2 - I^2} + (w + I) \cos^2\Theta -$$
$$- I - I_1 - 1 = 0 \qquad\qquad (13)$$

To determine, when u as a function of cosΘ has a maximum,

we might write out the differential of the first member for in-crements du, $d\cos\Theta$, and then put their quotient $= 0$. But that amounts to putting the factor of $d\cos\Theta$ equal to zero. Thus the maxima are determined by

$$- 2u \sqrt{w^2 - I^2} + 2(w + I) \cos\Theta = 0.$$

We have the legitimate presumption that the maximum values of u are 1; hence

$$\cos\Theta = \sqrt{\frac{w - I}{w + I}} \qquad (14)$$

gives the two directions, symmetrical with respect to the + 1-direction. It is easy to check that (13) actually has the root u = 1 for this value of $\cos\Theta$. The extreme cases mentioned before are correctly encompassed by $w = I$ and $I = 0$, respectively. –

It is hardly necessary to point out that all these deviations from normal behaviour, i. e. from u = 1, are presumably minute, since the unities in which our fields B, E, etc. are measured are presumably extremely outsized. Even so, the consequences of non-linearity have some interest by principle. Easily the quaintest event is, that in an extremely outsized background field the displacement (12) can become bigger than the 1-half-axis of the ellipsoid, which then excludes the origin, so that within a certain cone of wave-normals no positive u is available, and wave fronts cannot proceed in those directions. The con-dition for this freak to occur is, from (12) and (11),

$$\sqrt{w^2 - I^2} \geqq \sqrt{1 + 2I_1 - I_2^{\ 2}}$$

or
$$w^2 \geqq 1 + 2I_1 - I_2^{\ 2} + I^2 = (1 + I_1)^2$$

$$w \geqq 1 + I_1$$

or
$$\tfrac{1}{2}(B^2 + E^2) \geqq 1 + \tfrac{1}{2}(B^2 - E^2)$$

or
$$E^2 \geqq 1. \tag{15}$$

But can that be?   — Well, the only irrescindible requirement
in Born's electrodynamics is

$$A^2 = \frac{1 + I_1 - I}{1 + I_1 + I} = \frac{(1 + I_1)^2 - I^2}{(1 + I_1 + I)^2} \geqq 0.$$

Thus

$$(1 + I_1)^2 - I^2 = 1 + 2I_1 - I_2^{\ 2} \geqq 0$$

or

$$1 + B^2 - E^2 - (BE)^2 \geqq 0$$

or

$$E^2 \leqq \frac{1 + B^2}{1 + B^2 \cos^2 \alpha}, \tag{16}$$

where α is the angle between B and E.   Hence E^2 is well
allowed to surpass 1, provided that $B^2 \neq 0$ and $\cos^2 \alpha \neq 1.$

One might be inclined to brush aside this freak (as I
called it) because it requires field strengths that cannot ever
be reached.   Far from this, given any non-vanishing (E,B)-field

whatsoever, you can, in a suitable Lorentz frame, make E^2 sur-
pass unity.   This is almost self-evident, since the independ-
ent invariants are

$$\tfrac{1}{2}(B^2 - E^2), \qquad (BE).$$

Thus $|B|$ and $|E|$ may be increased indefinitely, in step with
each other, provided that $\cos\alpha$ approaches to zero, so as to
keep also the second invariant invariant.