ON SYSTEMS OF COORDINATE TRANSFORMATIONS AND WAVE MECHANICS

TONY LYONS

Abstract. The principles of special and general relativity are used to describe the relative motion of observers in terms of the coordinate transformations between the rest frames associated with each observer. It is demonstrated that a consistent application of these principles requires one to ascribe to each observer an associated wave amplitude which governs the probability events within one frame may be observed in another, leading the quantisation of energy levels under appropriate conditions.

1. Introduction

While it is always our aim to write the laws of mechanics in a general form independent of any particular coordinate system, we must be aware that all mechanical quantities are themselves always defined with reference to some coordinate system. The simplest mechanical system of all, that of a uniformly moving point particle will have the character of its motion altered when viewed from coordinate systems which are themselves in a state of relative motion. However, associated with every such particle there is a preferred coordinate system we refer to as the *rest frame* of that particle, wherein the particle remains at some fixed spatial location for all times.

Central to our understanding of mechanical laws is the principle of relativity which essentially asserts that the laws of mechanics are valid in all frames of reference, a principle which has been confirmed in a wealth of contexts. In the current work we shall use this principle as the motivation of our starting point, that the space-time coordinates used by any particular observer are valid for the description of mechanical process. In particular, it means that to with every body which we ourselves would describe as a moving body, there is associated an reference frame in which that body is always at rest, and which with full validity may be used as a frame of reference for the description of mechanical phenomena. As we will see in the following, this observation will allow us to interpret these various states of motion between bodies, in terms of the coordinate transformations between their associated rest frames only.

2. Systems of coordinate transformations

Throughout the term "observer" is used in the broad sense of an object that may serve as a point of reference to describe the relative motion of other objects. Consider two such observers O and O', where O is at rest with reference to the coordinate frame Σ, while the observer O' is at rest in relation to the coordinate frame Σ'. Without loss of generality the observer O is assigned the space-time location $(t,0,0,0) \in \Sigma$ for all $t \in \mathbb{R}$, while the observer O' is assigned the space-time location $(t',0,0,0) \in \Sigma'$ for all $t' \in \mathbb{R}$. On the other hand, the observer O' will occupy some location $(t,x,y,z) \in \Sigma$ where the x, y and z coordinates of O' may be functions of t, depending on the state of relative motion between the two observers. In compliance with the principle of relativity, this means the observer O will occupy some location (t',x',y',z') whose spatial coordinates x', y' and z' may also be functions of t', depending on the relative state of motion between the pair.

2.1. Uniform relative motion and coordinate transformations.

To simplify the description of a state of relative motion between O and O' it is assumed that the (x,y,z)-axes of the frame Σ align with the respective (x',y',z')-axes of the frame Σ'. Furthermore the location of O' with reference to Σ will take the form $(t,x,0,0)$ and similarly the location of O with respect to Σ' will take the form $(t',x',0,0)$. As such, when considering the location of observer O' within Σ it is only necessary to refer to its (t,x)-coordinates, and similarly the location of observer O with reference to Σ' will be described by its (t',x')-coordinates.

A state of uniform relative motion between O and O' will appear in the reference frame Σ as the observer O changing its space-time location according to $(t,0) \to (t+dt,0)$ while the observer O' will change its coordinate location according to $(t,x) \to (t+dt,x+dx)$. On the other hand, as viewed within the reference frame Σ', the observer O' changes its space-time coordinate according to $(t',0) \to (t'+dt',0)$, while the space-time location of O within Σ' will vary according to $(t',x') \to (t'+dt',x'+dx')$. A state of uniform motion will be taken to mean that the spatial displacement of O' within Σ, namely dx, is related to the temporal change dt by a *constant* factor v giving $dx = vdt$. Equivalently, within the frame Σ' the displacement of O, namely dx', will be proportional to the temporal change dt' according to $dx' = -vdt'$. The coordinate changes dx, dx', dt and dt' may be finite and we do not as yet restrict ourselves to infinitesimal coordinate changes.

Under this scheme it is clear that to every space-time location $(t,x) \in \Sigma$ there are corresponding coordinates $(t',x') \in \Sigma'$, and as such we may introduce a pair of maps $\mathbf{\Phi}'$ and $\mathbf{\Phi}$ such that

$$\begin{aligned} &\mathbf{\Phi}' : \Sigma \to \Sigma' \\ &\mathbf{\Phi}' : \begin{pmatrix} x \\ t \end{pmatrix} \mapsto \begin{pmatrix} \mathrm{T}'(t,x) \\ \mathrm{X}'(t,x) \end{pmatrix} = \begin{pmatrix} t' \\ x' \end{pmatrix} \end{aligned}$$

with the inverse given by

$$\begin{aligned} &\mathbf{\Phi} : \Sigma' \to \Sigma \\ &\mathbf{\Phi} : \begin{pmatrix} t' \\ x' \end{pmatrix} \mapsto \begin{pmatrix} \mathrm{T}(t',x') \\ \mathrm{X}(t',x') \end{pmatrix} = \begin{pmatrix} t \\ x \end{pmatrix}. \end{aligned}$$

Remark 1. Importantly we observer that the spatial locations within either frame Σ or Σ' are not necessarily a priori separate locations. To emphasise this we consider the observer O' within the frame Σ, which may undergo a displacement from $x \to x + dx$ in some time interval dt. However, within the frame Σ' no such separation of spatial locations is realised, instead the observer O' perceives this as only an increment of the time coordinate $t' \to t' + dt'$. In line with the principle of special relativity, the apparent spatial separation of the points x and $x + dx$ within Σ is realised as such only in that frame. Thus we do not ascribe any fundamental significance to spatial separations as observed within any reference frame.

2.2. **The Lorentz Transformations.** The derivation of the Lorentz transformation between coordinate systems in uniform relative motion relies on two significant conditions:

1) Linearity of the maps $\mathbf{\Phi}'(t, x)$ and $\mathbf{\Phi}(t', x')$ with respect to both of their arguments

2) A universally agreed upon velocity c (with which electromagnetic waves are known to propagate) within any frame of reference

(see [3] for further discussion regarding the need for linearity of the Lorentz transformations). Under these conditions, the coordinate maps introduced in Section 2.1 are also seen to correspond to the well studied Lorentz transformations. To establish these transformations in the current framework, we consider a pair of events as observed within Σ with space-time coordinates (t, x) and $(t+dt, x+dx)$, while those same events as observed within the frame Σ' occur at space-time coordinates (t', x') and $(t' + dt', x' + dx')$, where again, despite notation, the separations dt, dx, dt' and dx' may be finite (as opposed to infinitesimals).

Given the linearity of $\mathbf{\Phi}'$ we require it to be a coordinate transformation of the form

$$\begin{aligned} t' &= \mathrm{T}'(t,x) = kt + lx \\ x' &= \mathrm{X}'(t,x) = rt + sx, \end{aligned} \tag{1}$$

up to additive constants, where k, l, r and s are all independent of (t, x) (and (t', x')). Owing to the linearity of $\mathbf{\Phi}'$ it follows that the coordinate separation of the events within Σ' are realised as

$$dt' = \mathrm{T}'(t + dt, x + dx) - \mathrm{T}'(t, x) = kdt + ldx$$
$$dx' = \mathrm{X}'(t + dt, x + dx) - \mathrm{X}'(t, x) = rdt + sdx.$$

Thus the velocity along the trajectory connecting the events at (t', x') and $(t' + dt', x' + dx')$, as observed within Σ', is given by

$$\frac{dx'}{dt'} = \frac{r + s\frac{dx}{dt}}{k + l\frac{dx}{dt}}$$

where $\frac{dx}{dt}$ is the velocity along the trajectory connecting the events as observed within Σ.

The condition **2)** above means that if the velocity along the world line within Σ is $\frac{dx}{dt} = c$, then the velocity along the same world-line within Σ' is also $\frac{dx'}{dt'} = c$, and so we observe

$$c = \frac{r + sc}{k + lc}. \tag{2}$$

On the other hand, we assume the observer O' moves with a velocity v within Σ, in which case a world-line with velocity $\frac{dx}{dt} = v$ within Σ will appear with a fixed x'-coordinate within Σ', and so we deduce

$$0 = \frac{dx'}{dt'} = \frac{r + sv}{k + lv} \Rightarrow r = -sv. \tag{3}$$

Moreover, since observer O moves with velocity $-v$ relative to Σ', it follows that a world-line with velocity $\frac{dx}{dt} = 0$ in Σ will have velocity $\frac{dx'}{dt'} = -v$ within Σ', meaning

$$-v = \frac{r}{k} \Rightarrow r = -kv. \tag{4}$$

Thus equations (3)-(4) ensure

$$k = s,$$

while equation (2) now means

$$l = -\frac{kv}{c^2}.$$

To establish the form of k we write the coordinate transformation (1) in matrix form to give:

$$\mathbf{\Phi}'(t, x) = \begin{bmatrix} \mathrm{T}'(t, x) \\ \mathrm{X}'(t, x) \end{bmatrix} = \begin{bmatrix} t' \\ x' \end{bmatrix} = \begin{bmatrix} k & l \\ r & s \end{bmatrix} \begin{bmatrix} t \\ x \end{bmatrix}$$

Inverting the square matrix in this transformation, we find

$$\begin{bmatrix} k & l \\ r & s \end{bmatrix}^{-1} = \frac{1}{J} \begin{bmatrix} s & -l \\ -r & k \end{bmatrix} \qquad J = ks - rl,$$

and so the inverse coordinate transformation $\mathbf{\Phi} : \Sigma' \to \Sigma$ is given by

$$\mathbf{\Phi}(t',x') = \begin{bmatrix} \mathrm{T}(t',x') \\ \mathrm{X}(t',x') \end{bmatrix} = \begin{bmatrix} t \\ x \end{bmatrix} = \frac{1}{J}\begin{bmatrix} s & -l \\ -r & k \end{bmatrix}\begin{bmatrix} t' \\ x' \end{bmatrix}.$$

It follows that the coordinates (t,x) and (t',x') may also be related according to

$$\begin{aligned} t &= \frac{1}{J}\left(lt' - sx'\right) = \frac{k}{J}(t' + \frac{v}{c^2}x') \\ x &= \frac{1}{J}\left(-kt' + rx'\right) = \frac{k}{J}(x' + vt'). \end{aligned} \tag{5}$$

However, a consequence of the principle of relativity is that the transformations are symmetric, in that $\mathbf{\Phi}$ may be obtained from $\mathbf{\Phi}'$ (and vice-verse) by simply changing the sign of the velocity and swapping the roles of (t,x) and (t',x'). It follows from equation (1) and the principle of special relativity that the coordinates (t,x) and (t',x') may also be related according to

$$\begin{aligned} t &= k(t' + \frac{v}{c^2}x') \\ x &= k(x' + vt'), \end{aligned} \tag{6}$$

and upon comparing (5) and (6) it becomes clear that $J = 1$ and so

$$k = \pm\frac{1}{\sqrt{1 - \frac{v^2}{c^2}}},$$

with $k > 0$ corresponding to the restricted Lorentz group, that is to say the space of Lorentz transformations continuously connected to the identity transformation.

Remark 2. We note that this derivation of the Lorentz transformation does not require O and O' to be both inertial observers. Indeed, the derivation above may allow for both O and O' have the same acceleration relative to a third observer, it is only necessary that the relative motion between the pair is uniform.

The reference frames Σ with coordinates (t,x,y,z) and Σ' with coordinates (t',x',y',z'), whose alignment and relative velocity we have already stipulated, are easily seen to be related according to

$$\begin{aligned} (t',x',y',z') &= \mathbf{\Phi}'(t,x,y,z) \\ \mathrm{T}'(t,x,y,z) &= t' = \frac{\left(t-\frac{v}{c^2}x\right)}{\sqrt{1-\frac{v^2}{c^2}}} \\ \mathrm{X}'(t,x,y,z) &= x' = \frac{(x-vt)}{\sqrt{1-\frac{v^2}{c^2}}} \\ \mathrm{Y}'(t,x,y,z) &= y' = y \\ \mathrm{Z}'(t,x,y,z) &= z' = z, \end{aligned}$$

as expected, and whose inverse is given by

$$\begin{aligned} (t,x,y,z) &= \mathbf{\Phi}(t',x',y',z') \\ \mathrm{T}(t',x',y',z') &= t = \frac{\left(t'+\frac{v}{c^2}x'\right)}{\sqrt{1-\frac{v^2}{c^2}}} \\ \mathrm{X}(t',x',y',z') &= x = \frac{(x'+vt')}{\sqrt{1-\frac{v^2}{c^2}}} \\ \mathrm{Y}(t',x',y',z') &= y = y' \\ \mathrm{Z}(t',x',y',z') &= z = z'. \end{aligned}$$

Let $E'_0 = mc^2$ be the rest energy of observer O', that is to say its energy with respect to its rest frame Σ', in which case the Lorentz transformation may now be written according to

$$\begin{aligned} \mathrm{T}'(t,x,y,z) &= t' = \frac{1}{mc^2}\left(Et-px\right) \\ \mathrm{X}'(t,x,y,z) &= x' = \frac{1}{mc^2}\left(Ex-pc^2t\right) \\ \mathrm{Y}'(t,x,y,z) &= y' = y \\ \mathrm{Z}'(t,x,y,z) &= z' = z, \end{aligned} \tag{7}$$

where $(E,p,0,0) = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}\left(mc^2,mv,0,0\right)$ is the energy-momentum four-vector of the observer O' with reference to the frame Σ. As we can clearly see, the coordinate transformation $\mathbf{\Phi}'$ is characterised by the energy-momentum four vector of O' within the frame Σ.

2.3. **Relative acceleration.** When O and O' are moving relative to each other with non-constant velocity it is still the case that the observers have associated reference frames Σ and Σ' in which each is always at rest. In general it is still expected that the space time location in one reference frame has corresponding space time coordinates in the other, with maps from one frame to the other expected to satisfy $(\mathbf{\Phi}\circ\mathbf{\Phi}')(t,x) = (t,x)$ and likewise $(\mathbf{\Phi}'\circ\mathbf{\Phi})(t',x') = (t',x')$ for all $(t,x)\in\Sigma$ (resp. for all $(t',x')\in\Sigma'$).

At any instant of this relative motion these coordinate systems are require to satisfy the inverse function theorem

$$\begin{bmatrix} \mathrm{T}_{t'}(t',x') & \mathrm{X}_{t'}(t',x') \\ \mathrm{T}_{x'}(t',x') & \mathrm{X}_{x'}(t',x') \end{bmatrix} \begin{bmatrix} \mathrm{T}'_t(t,x) & \mathrm{X}'_t(t,x) \\ \mathrm{T}'_x(t,x) & \mathrm{X}'_x(t,x) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{8}$$

(cf.[8]) where subscripts denote partial differentiation with respect to the relevant variable. Multiplying by mc^2 this matrix equation becomes

$$\begin{bmatrix} \mathrm{T}_{t'}(t',x') & \mathrm{X}_{t'}(t',x') \\ \mathrm{T}_{x'}(t',x') & \mathrm{X}_{x'}(t',x') \end{bmatrix} \begin{bmatrix} E & -p \\ -pc^2 & E \end{bmatrix} = \begin{bmatrix} mc^2 & 0 \\ 0 & mc^2 \end{bmatrix},$$

having used equation (7) to write $\mathbf{\Phi}(t,x)$ in terms of (E,p) the energy momentum of O_0 with reference to Σ. We may write the matrices as such since at any instant the observer O' moves with some energy-momentum relative to the frame Σ and the Lorentz transformation between Σ and Σ' should be valid at this the location of O' at least locally.

It follows at once that

$$\begin{bmatrix} \mathrm{T}_{t'}(t',x') & \mathrm{X}_{t'}(t',x') \\ \mathrm{T}_{x'}(t',x') & \mathrm{X}_{x'}(t',x') \end{bmatrix} = \frac{1}{mc^2} \begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix},$$

and as such equation (8) may also be written according to

$$\begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix} \begin{bmatrix} \mathrm{T}'_t(t,x) & \mathrm{X}'_t(t,x) \\ \mathrm{T}'_x(t,x) & \mathrm{X}'_x(t,x) \end{bmatrix} = \begin{bmatrix} mc^2 & 0 \\ 0 & mc^2 \end{bmatrix}.$$

Differentiation with respect to t yields

$$\left(\partial_t \begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix}\right) \begin{bmatrix} \mathrm{T}'_t & \mathrm{X}'_t \\ \mathrm{T}'_x & \mathrm{X}'_x \end{bmatrix} = -\begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix} \partial_t \begin{bmatrix} \mathrm{T}'_t & \mathrm{X}'_t \\ \mathrm{T}'_x & \mathrm{X}'_x \end{bmatrix}$$

from which it follows

$$mc^2 \begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix}^{-1} \left(\partial_t \begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix}\right) \begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix}^{-1} = -\begin{bmatrix} \Gamma^0_{00} & \Gamma^1_{00} \\ \Gamma^0_{10} & \Gamma^1_{10} \end{bmatrix}$$

having used $\begin{bmatrix} E & p \\ pc^2 & E \end{bmatrix}^{-1} = \frac{1}{m^2c^4} \begin{bmatrix} E & -p \\ -pc^2 & E \end{bmatrix}$ and introduced the Christoffel coefficients $\Gamma^0_{00} = \mathrm{T}'_{tt}$, $\Gamma^1_{00} = \mathrm{X}'_{tt}$, $\Gamma^0_{10} = \mathrm{T}'_{xt}$ and $\Gamma^1_{10} = \xi_{xt}$. Given $\mathbf{A}\mathbf{A}^{-1}$ is the constant identity matrix for any matrix $\mathbf{A}$ it also follows that

$$\partial_\mu \mathbf{A}^{-1} = -\mathbf{A}^{-1}(\partial_\mu \mathbf{A})\mathbf{A}^{-1},$$

and so we deduce

$$\partial_t \begin{bmatrix} E & -p \\ -pc^2 & E \end{bmatrix} = mc^2 \begin{bmatrix} \Gamma^0_{00} & \Gamma^1_{00} \\ \Gamma^0_{10} & \Gamma^1_{10} \end{bmatrix}$$

In a similar manner, it is easily shown that

$$\partial_x \begin{bmatrix} E & -p \\ -pc^2 & E \end{bmatrix} = mc^2 \begin{bmatrix} \Gamma^0_{01} & \Gamma^1_{01} \\ \Gamma^0_{11} & \Gamma^1_{11} \end{bmatrix}$$

with $\Gamma^0_{01} = \mathrm{T}'_{tx}$, $\Gamma^1_{01} = \mathrm{X}'_{tx}$, $\Gamma^0_{11} = \mathrm{T}'_{xx}$ and $\Gamma^1_{11} = \mathrm{X}'_{xx}$. Since it is always possible to introduce a coordinate transformation such that the Christoffel coefficients vanish locally, it is always possible to ensure the force $E_x = p_t$ experienced by any observer can be eliminated, however it means the force must appear to act on another observer in a state of relative acceleration at another location.

2.3.1. *Relative acceleration in a potential field.* We consider two observers O_a and O_b in a state of relative acceleration, with Σ the rest frame of O_a and Σ' the rest frame of O_b. Associated with the observer O_a is its mass m_a and similarly O_b has mass m_b. The observers are assumed to interact with one another via a potential field $(\phi(t, x), A(t, x))$, where the conserved energy of O' with reference to the frame Σ is given by

$$\mathcal{E}(t, x_b) = E_b + \phi(t, x_b)$$

while the associated generalised momentum(cf. [4] p. 48) is given by

$$\mathcal{P}(t, x_b) = p_b + A(t, x_b).$$

The energy-momentum of O with reference to Σ are

$$\mathcal{E}(t, x_a) = m_a c^4 + \phi(t, x_a) = m_a c^2 \quad \forall t \in \mathbb{R},$$

and

$$\mathcal{P}(t, x_a) = A(t, x_a) = 0 \quad \forall t \in \mathbb{R},$$

with $E_{a,b} = \sqrt{m^2_{a,b}c^4 + p^2_{a,b}c^2}$.

The potential field $(\phi(t, x), A(t, x))$ is required to transform four vector under the Lorentz transformation since $(\mathcal{E}(t, x), \mathcal{P}(t, x))$ is also required to change as a four vector. It follows that

$$\begin{aligned} \begin{bmatrix} \mathcal{E}'(t'_b, x'_b) \\ \mathcal{P}'(t'_b, x'_b) \end{bmatrix} &= \frac{1}{m_b c^2} \begin{bmatrix} E_b & -p_b c^2 \\ -p_b & E_b \end{bmatrix} \begin{bmatrix} \mathcal{E}(\mathrm{T}(t'_b, x'_b), \mathrm{X}(t'_b, x'_b)) \\ \mathcal{P}(\mathrm{T}(t'_b, x'_b), \mathrm{X}(t'_b, x'_b)) \end{bmatrix} \\ &= \begin{bmatrix} E_b & p_b c^2 \\ p_b & E_b \end{bmatrix} \begin{bmatrix} \mathcal{E}(\mathrm{T}(t'_b, x'_b), \mathrm{X}(t'_b, x'_b)) \\ \mathcal{P}(\mathrm{T}(t'_b, x'_b), \mathrm{X}(t'_b, x'_b)) \end{bmatrix} \\ &= \begin{bmatrix} m_b c^2 \\ 0 \end{bmatrix}, \end{aligned} \tag{9}$$

since $(\mathcal{E}'(t'_b, x'_b), \mathcal{P}'(t'_b, x'_b))$ is the total energy-momentum vector of O_b in its rest frame Σ' which is always $(m_b c^2, 0)$ by definition of Σ'. Using

$\mathcal{E}(\mathrm{T}(t'_b, x'_b), \mathrm{X}(t'_b, x'_b)) = E_b + \phi(t, x_b)$ and $\mathcal{P}(\mathrm{T}(t'_b, x'_b), \mathrm{X}(t'_b, x'_b)) = p_b + A(t, x_b)$ in the first line of (9) we have

$$\frac{1}{m_b c^2}\begin{bmatrix} E_b & p_b c^2 \\ p_b & E_b \end{bmatrix}\begin{bmatrix} \phi(t, x_b) \\ A(t, x_b) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Differentiating, it follows that

$$\begin{bmatrix} \phi_b \partial_\mu E_b + c^2 A_b \partial_\mu p_b + E_b(\partial_\mu \phi_b) + p_b c^2 \partial_\mu A_b \\ \phi_b \partial_\mu p_b + +A_b \partial_\mu E_b + p_b \partial_\mu \phi_b + A_b \partial_\mu A_b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where $\mu \in \{t, x\}$ and we use the abbreviations $\phi_b := \phi(t, x_b)$ and $\partial_\mu \phi_b := (\partial_\mu \phi)(t, x_b)$ with similiar definitions in operation for A_b and $\partial_\mu A_b$ and . It is immediately clear that

$$\begin{bmatrix} \partial_\mu E_b \\ \partial_\mu p_b \end{bmatrix} = \frac{1}{c^2 A_b^2 - \phi_b^2}\begin{bmatrix} \phi_b & -c^2 A_b \\ -A_b & \phi_b \end{bmatrix}\begin{bmatrix} E_b & p_b c^2 \\ p_b & E_b \end{bmatrix}\begin{bmatrix} \partial_\mu \phi_b \\ \partial_\mu A_b \end{bmatrix}.$$

3. Particles and waves

Thus far we have considered observers O and O' as strictly particle like, in that it is assumed they occupy a unique spatial location x (or x'_0) at any fixed moment t (or t') within the frame Σ (or Σ'). Moreover, it has been assumed that there is a unique map $\mathbf{\Phi'}$ from Σ to Σ', and vice-versa. In Figure 1 we illustrate a simple container comprised of two stationary walls located at x_a and x_b in the reference frame Σ. We suppose a particle of rest mass m is located between these walls and moving with some velocity relative to them. We further assume there is no interaction between this particle and the walls of the container, other than a possible collision between the particle and either wall. Such a collision will in turn set the walls of the container in motion, with the relative velocity of the walls allowing us to deduce the momentum change of the particle upon impact.

Crucially, we insist that the motion of the walls is defined in relative terms only, that is to say, the wall at x_a moves relative to the wall at x_b if and only if the wall at x_b moves relative to the wall at x_a. Let Σ be the rest-frame of the left-plate (which we call observer O) whose location therein we denote (x_a, t) for all $t \in \mathbb{R}$, while the initial location of the right plate (which we call observer O') is $x = x_b$. At some time $t = t_1$ the right plate begins to move and is observed at a new location $x = x_b + dx$ at a later time $t_1 + dt$, as illustrated in the top panel of Figure 1. Conversely in the frame Σ' the plate at $x' = x_a$ begins to move at $t' = t_1$ and at a later time $t' = t_1 + dt'$ this plate is located at $x_a + dx'$, as illustrated in the lower panel of Figure 1.

However, this appears to contradict our usual experience of particle motion. Our usual experience would indicate that a collision between the particle and the wall at x_a will cause this wall to accelerate and thus move relative to the wall at x_b, and conversely, an impact between the particle and the wall at x_b will cause this wall to accelerate and

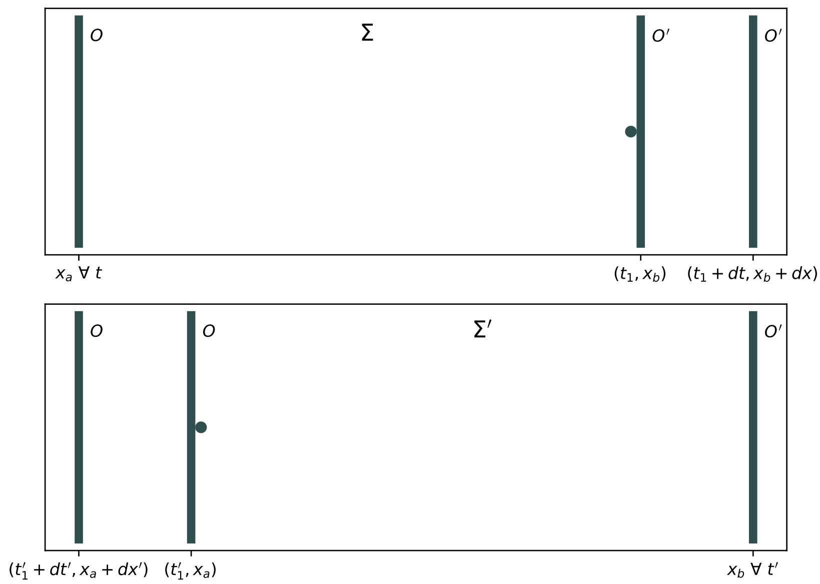

Figure 1. The relative acceleration of the plates in the rest-frame of the left plate Σ (top panel) and in the rest-frame of the right plate Σ' (bottom panel)

thus start moving relative to the wall at x_a. While we would at this point insist it is the acceleration of the impacted wall which breaks the symmetry of the problem, *it is not clear that the acceleration of the walls in this system can be defined in anything but relative terms either* (see [5] for further discussion on the principle of general relativity whereby all reference frames are considered equally valid for the description of physical processes). Consequently as to which wall is indeed moving appears to depend upon the reference frame used, and as such the actual location of the impact between the particle and the container is not defined in absolute terms. In such an experimental configuration the location of the particle is not in fact well defined and is only ever meaningfully defined up to a precision $l = x_b - x_a$.

3.1. **The Hamilton-Jacobi equation.** Consider now a particle in motion between the plates when these plates are a rest relative to each other. This particle may also serve as an observer we call O_0 and has an associated rest frame Σ_0 wherein its own space-time location has coordinates $(\tau, 0)$ for $\tau \in \mathbb{R}$. The locations of the observers O and O' will be of the form (τ, ξ_a) and (τ, ξ_b) where ξ_a and ξ_b are functions of τ depending on the relative state of motion between O_0 and the observers O and O'. The map from the coordinate frame Σ to the frame Σ_0 and its inverse are denoted by

$$\mathbf{\Psi_0}(t,x) = \begin{bmatrix}\tau(t,x)\\ \xi(t,x)\end{bmatrix} = \begin{bmatrix}\tau\\ \xi\end{bmatrix} \quad \mathbf{\Psi}(\tau,\xi) = \begin{bmatrix}\mathrm{T}(\tau,\xi)\\ \mathrm{X}(\tau,\xi)\end{bmatrix} = \begin{bmatrix}t\\ x\end{bmatrix}$$

It is noted that Σ and Σ' are the same coordinate system up to a simple translation while O and O' are at rest with respect to each other.

The quantum mechanical description of matter due to Schrödinger relies on the description of a material particle in terms of the Hamilton-Jacobi equation for the system, usually presented in the form

$$\frac{\partial S}{\partial t} + H\left(q, \frac{\partial S}{\partial q}\right) = 0,$$

The functional

$$S = \int_{t_0}^{t} L\left[q, \dot{q}\right] dt$$

is the classical action associated with the particle motion, $L\left[q, \dot{q}\right]$ being the associated Lagrangian and $H(q,p)$ the corresponding Hamiltonian obtained from the Legendre transformation

$$H(q,p) = p\dot{q} - L\left[q, \dot{q}\right], \quad p = \frac{\partial L}{\partial \dot{q}},$$

(see [3] for a comprehensive discussion of the Hamilton-Jacobi equations in classical mechanics).

In the current scenario the generalised coordinate q of the observer O_0 corresponds to its spatial coordinate x (or ξ) within the reference frame Σ (or Σ_0). In particular, the classical action of the observer with respect to its own rest frame Σ_0 is simply

$$S = -\int_{\tau_0}^{\tau} mc^2 d\tau, \tag{10}$$

this action being invariant under the group of Lorentz transformations. In terms of the Hamilton-Jacobi formalism the energy-momentum of O_0 with reference to the frame Σ are given by

$$E = -\frac{\partial S}{\partial t} \quad p = \frac{\partial S}{\partial x},$$

in line with the transformation (7).

3.2. **The wave equation.** In [9, 10] a substitution of the form $S = K\log\psi$ was used to demonstrate that a Keplerian system was describable in terms of a *wave function* ψ where the quantisation of energy levels in the system was understood in terms analogous to the discrete modes of a vibrating string. In some sense this result is a formalisation of de Broglie's original proposition that a complete description of particle motion should be understood in terms of associated wave properties (cf. [1, 2]). In the later work [7] it was demonstrated that a particle following a trajectory governed by a classical action of the form (10) may be described in terms of a wave-function which shares many similarities with de Broglie's.

The associated wave equation is derived from the relativistic energy-momentum relation

$$E^2 - p^2c^2 = m^2c^4$$

and the condition ensuring conservation of energy-momentum along the trajectory, namely

$$\frac{1}{c^2}\frac{\partial E}{\partial t} - \frac{\partial p}{\partial x} = 0.$$

These conditions written using the Hamilton-Jacobi formalism are given by

$$\frac{1}{c^2}\left(\frac{\partial S}{\partial t}\right)^2 - \left(\frac{\partial S}{\partial x}\right)^2 = m^2c^2 \tag{11a}$$

$$\frac{1}{c^2}\frac{\partial^2 S}{\partial t^2} - \frac{\partial^2 S}{\partial x^2} = 0, \tag{11b}$$

respectively. Following a procedure similar to that adopted in [9] we propose $S = K\ln\psi$ where K is constant, which upon substitution into (11a)–(11b) yields

$$\frac{1}{c^2}\frac{\partial^2 \psi}{\partial t^2} - \frac{\partial^2 \psi}{\partial x^2} - \frac{m^2c^2}{K^2}\psi = 0. \tag{12}$$

Meanwhile the function ψ may be written in terms of the action integral according to

$$\psi(t,x) = \exp\left\{-\frac{1}{K}\int_{\tau(t_0,x_0)}^{\tau(t,x)} mc^2 d\tau\right\}.$$

In [6, 7] this expression for ψ is interpreted as *the probability amplitude observer* O_0 *is located somewhere along the classical trajectory with end points* (t_0, x_0) *and* (t, x).

Given the map $\tau(t,x)$ (cf. equation (7)), we note that at any fixed instant t there is a spread in the *possible* values of τ associated with the observer O_0 given by

$$\Delta\tau = \tau(t,x_a) - \tau(t,x_b) = \frac{pl}{mc^2}, \tag{13}$$

depending on where O_0 is with respect to Σ. If the plates of the apparatus begin to move apart at some instant t then it is required that the probability amplitude at x_a and x_b should be of equal magnitude which requires $|\psi(t,x_a)| = |\psi(t,x_a)|$, thus ensuring it is equally likely the particle is at x_a or x_b at the moment of impact. Thus we deduce

$$\pm 1 = \exp\left\{-\frac{1}{K}\int_{\tau(t,x_a)}^{\tau(t,x_b)} mc^2 d\tau\right\} = e^{-\frac{1}{K}mc^2\Delta\tau},$$

which simply follows from $\pm\psi(t, x_a) = \psi(t, x_b)$ having divided through by $\psi(t, x_a)$. Having insisted O_0 should have non-zero velocity relative to the frame Σ it follows that $\Delta\tau \neq 0$ (cf. equation (13)) and so it follows that $\frac{mc^2\Delta\tau}{K} = n\pi i$ for $n \in \mathbb{Z}$ (since $e^x > 0$ and $e^x = 1 \Leftrightarrow x = 0$ for $x \in \mathbb{R}$). As such we set $K = \frac{\hbar}{i}$ with $\hbar$ the reduced Planck constant, and equation (12) becomes

$$\frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} - \frac{\partial^2\psi}{\partial x^2} + \frac{m^2c^2}{\hbar^2}\psi = 0, \tag{14}$$

the familiar Klein-Gordon equation.

The basis of the solution space for equation (14) is of the form $e^{\pm i(\omega t - kx)}$ where $\hbar\omega = E$ and $\hbar k = p$ with $E^2 - p^2c^2 = m^2c^4$, while a general solution $\psi(t, x)$ may be written as a superposition of such basis solutions. Consider now the simplest such superposition

$$\psi(t, x) = a(k)e^{i(\omega t - kx)} + a(-k)e^{i(\omega t + kx)}$$

where O_0 is in a superposition of possible momentum states $\pm p$ with respect to the frame Σ as illustrated in Figure 2. It is important to

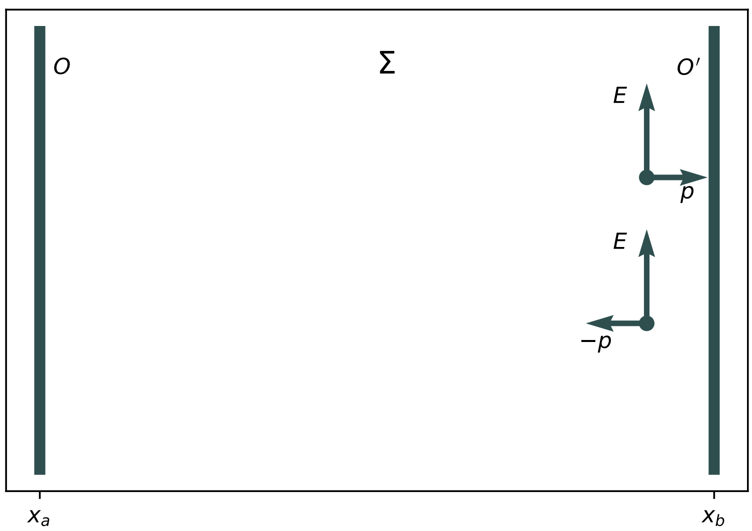

FIGURE 2. The possible momentum states of O_0 as observed in the reference frame Σ, the rest frame of O

note that in Figure 2 the observer O_0 is not in collision with the plate at x_b in the classical sense as this would mean a transfer of energy-momentum from O_0 to O', which would in turn require O_0 to undergo a reduction in energy after the collision.

In the specific case where $a(k) = -a(-k)$ we find

$$\psi(t, x) = -2ia(k)\sin(kx)e^{i\omega t},$$

and imposing $\int_{x_a}^{x_b} |\psi(t,x)|^2\, dx = 1$, we find

$$1 = |a(k)|^2 \left[2l + \cos(2kx_b) - \cos(2kx_a)\right]$$

This must be true for any values $x_b = x_a + l$ since the coordinates are simply labels, and so for the choice $-x_a = x_b = \frac{l}{2}$ it follows that

$$1 = 2l\, |a(k)|^2 ,$$

which must also be true for any $x_b = x_a + l$, meaning

$$2kx_b = 2kx_a + 2n\pi \quad n \in \mathbb{Z}.$$

It follows at once that

$$k = \frac{n\pi}{l},$$

as is familiar from the non-relativistic particle in a box and that such a superposition generally takes the form

$$\psi_n(t,x) = \sqrt{\frac{2}{l}} \sin\left(\frac{n\pi x}{l}\right) e^{i(\omega_n t - \phi)}, \tag{15}$$

with $\omega_n = \sqrt{\frac{m^2c^4}{\hbar^2} + \frac{n^2\pi^2c^2}{l^2}}$ the angular frequency associated with the n^{th}-energy level $E_n = \hbar\omega_n$ and where ϕ is an arbitrary real constant phase.

3.3. **Interpretation of the wave function.** It is clear that

$$\exp\left\{\frac{i}{\hbar}\int_{\tau(t,x_a)}^{\tau(t,x_b)} mc^2 d\tau\right\}$$

cannot represent the probability amplitude a particle is somewhere along the trajectory connecting (t, x_a) and (t, x_b) since the particle cannot travel between these two locations without violating causality. Thus we abandon the interpretation that $\exp\left\{\frac{i}{\hbar}\int_{\tau(t_0,x_0)}^{\tau(t,x)} mc^2 d\tau\right\}$ is the probability amplitude the particle is somewhere along a trajectory connecting (t_0, x_0) and (t, x) even if the two space-time locations are causally connected.

Instead we recall what the map $\tau(t, x)$ actually represents classically: it is the time coordinate ascribed to an even at (t, x) according to the observer O_0 who is at rest with respect to the coordinate frame Σ_0. Thus $\exp\left\{\frac{i}{\hbar}\int_{\tau(t_0,x_0)}^{\tau(t,x)} mc^2 d\tau\right\}$ is the probability amplitude that events as observed in Σ at (t_0, x_0) and (t, x) will appear at separate times in Σ_0 the time difference being $\tau(t, x) - \tau(t_0, x_0)$. Now owing to the linearity of the map $\tau(t, x)$, fixing the values (t_0, x_0) and (t, x) means a specific value for the difference $\tau(t, x) - \tau(t_0, x_0)$ ensures unique values for (E, p). As such the probability amplitude $\tau(t_0, x_0) - \tau(t, x)$ has a certain value is equivalent to the probability amplitude the observer O_0

has an appropriate energy-momentum (E, p) with respect to the frame Σ to ensure this difference in τ-values.

Consider now the observer O_0 whose motion with respect to the coordinate system is characterised by the wave function $\psi_n(t, x)$ given in equation (15), and an event whose coordinates in Σ are of the form (t, x_b). Since $\psi_n(t, x)$ is the probability amplitude that the corresponding τ-coordinate of the event in the frame Σ_0 is $\tau(t, x)$, then the probability amplitude any event of the form (t, x_b) will have some τ-coordinate in Σ becomes zero. In particular there is a vanishing probability density $|\psi(t, x_b)|$ associated with a coordinate correspondence of the form $(t, x_b) \simeq (\tau, 0)$, where $\simeq$ is used to indicate these are space time coordinates for the same event as observed in different frames. This in turn means O_0 and O' never occupy the same space time location. Thus there is a vanishing probability density associated with a collision between the O_0 and O' when the coordinate map from Σ to Σ_0 is characterised by $\psi_n(t, x)$.

References

[1] L de Broglie, *Recherches sur la theorie des quanta*, Ann. Phys. **10** (1925), 22–128.

[2] L de Broglie and H T Flint, *An Introduction to the Study of Wave Mechanics*, Methuen & Company Limited, 1930.

[3] H Goldstein, C Poole, and J Safko, *Classical Mechanics*, American Association of Physics Teachers, 2002.

[4] L D Landau and E M Lifshitz, *The Classical Theory of Fields*, Course of Theoretical Physics, vol. 2, Butterworth-Heinemann, 2013.

[5] C. Møller, *The Theory of Relativity*, International Series of Monographs on Physics, Clarendon Press, 1972.

[6] L Motz, *Quantization and the classical Hamilton-Jacobi equation*, Phys. Rev. **126** (1962), 378.

[7] L Motz and A Selzer, *Quantum mechanics and the relativistic Hamilton-Jacobi equation*, Phys. Rev. **133** (1964), no. 6B, B1622.

[8] W Rudin, *Principles of Mathematical Analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill New York, 1976.

[9] E. Schrödinger, *Quantisierung als eigenwertproblem*, Ann. Phys. **384** (1926), 361–376.

[10] E Schrödinger, *Quantization as a problem of proper values (Part i)*, Collected papers on wave mechanics, vol. 302, American Mathematical Society, 2003.

Department of Computing & Mathematics, Waterford Institute of Technology, Waterford, Ireland.

Email address: tlyons@wit.ie