HAL open science

Energy-mass equivalence from Maxwell equations

Alejandro Perez, Salvatore Ribisi

To cite this version:

Alejandro Perez, Salvatore Ribisi. Energy-mass equivalence from Maxwell equations. American Journal of Physics, 2022, 90 (4), pp.305-313. 10.1119/10.0009156 . hal-03136696

HAL Id: hal-03136696

https://hal.science/hal-03136696v1

Submitted on 31 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Energy-mass equivalence from Maxwell equations

Alejandro Perez* and Salvatore Ribisi†
Aix Marseille Univ, Université de Toulon, CNRS, CPT, 13000 Marseille, France

Since the appearance of Einstein's paper *"On the Electrodynamics of Moving Bodies"* and the birth of special relativity, it is understood that the theory was basically coded within Maxwell's equations. The celebrated mass-energy equivalence relation, $E = mc^2$, is derived by Einstein using thought experiments involving the kinematics of the emission of light (electromagnetic energy) and the relativity principle. Text book derivations often follow paths similar to Einstein's, or the analysis of the kinematics of particle collisions interpreted from the perspective of different inertial frames. All the same, in such derivations the direct dynamical link with hypothetical fundamental fields describing matter (e.g. Maxwell theory or other) is overshadowed by the use of powerful symmetry arguments, kinematics, and the relativity principle.

Here we show that the formula can be derived directly form the dynamical equations of a massless matter model confined in a box (which can be thought of as a toy model of a composite particle). The only assumptions in the derivation are that the field equations hold and the energy-momentum tensor admits a universal interpretation in arbitrary coordinate systems. The mass-energy equivalence relation follows from the inertia or (taking the equivalence principle for granted) weight of confined field radiation. The present derivation offers an interesting pedagogical perspective on the formula providing a simple toy model on the origin of mass and a natural bridge to the foundations of general relativity.

PACS numbers:

arXiv:2101.11923v1 [gr-qc] 28 Jan 2021

I. INTRODUCTION

One of the striking results of special relativity [1] is the implication of an equivalence between the concepts of inertia and energy. In one of his founding papers [2] Einstein arrives at the postulate of mass-energy equivalence by showing that a body emitting an energy E via electromagnetic radiation will see its mass decreased by an amount E/c^2. Today textbooks give several different derivations. In the classic book [3], for instance, the equivalence is found by equating the force on a charged particle with its four-momentum variation. Another derivation is presented that uses the consistency with the relativity principle of the kinematics of colliding particles as seen from different inertial frames. Perhaps the simplest (yet the most formal) derivation corresponds to the one that starts from the geometric (relativistic) free particle action

$$S[x(t)] = -mc \int dt \sqrt{|g_{\mu\nu}\dot{x}^\mu \dot{x}^\nu|} \tag{1}$$

whose non relativistic limit justifies the non-relativistic Lagrangian $L = m\dot{x}^2/2$, and in literally two lines of Hamiltonian analysis produces the canonical Hamiltonian energy $E(v) = mc^2/(\sqrt{1 - v^2/c^2})$ with $E(0) = mc^2$.

All these derivations are important and insightful in their own way and remain perhaps the simplest path to the equivalence formula. However, none of these make the link between pure energy and mass dynamically explicit. They hide, in some sense, one very important aspect which is perhaps the central one in view of the necessary generalization of special relativity to include gravity in the general theory of relativity.

The derivation proposed here is complementary to the standard account with the added value of presenting an explicit link between dynamics of a massless matter model—here the electromagnetic field, or a massless scalar field—and mass of its energy when confined in an idealized box. The derivation uses strongly the notion of stress-energy-momentum tensor of matter and the covariant interpretation of its physical content. The exercise paves the way for the understanding of the energy momentum tensor as the source of gravity in general relativity, and serves also as an introduction to the mathematics that is central in the definition of the theory.

We will consider radiation confined in an idealised box and we will show that this radiation confers inertia—encoded in a mass $m = E/c^2$—to the box, where E is its energy content. The simplest model of radiation will be

*Electronic address: perez@cpt.univ-mrs.fr
†Electronic address: salvatore.ribisi@cpt.univ-mrs.fr

Maxwell electromagnetic fields which strengthens the idea that special relativity is entirely encoded in the properties of electromagnetism (this was indeed the perspective adopted by Einstein in his 1905 revolution). As our derivation basically relies entirely on coordinate independence of the field equations and the conservation of the stress-energy-momentum tensor of the matter fields, the result should be valid for any generally covariant matter model. As an additional example we show that the construction also works for a massless scalar field.

The idealized box confining the radiation can in turn be thought of as a poor's man model of a composite particle (such as a proton or a neutron): from modern computations [4] in quantum chromodynamics (QCD) we know that 99% of the mass of a proton comes from the energy of the confined (not by a box but by the non-linear strong interaction) gluon-quark radiation while only the remaining 1% is associated to the contribution of the rest mass of the quarks. The QCD confinement potential is replaced in our model by the box and the boundary conditions that require the fields to vanish at its walls [1].

Our derivation could not realistically have replaced the historical one because it strongly uses the physical interpretation of the stress-energy-momentum tensor (a notion that even when present in the literature on Maxwell fields, became only central after the development of relativity) and the general covariance of the relativistic field equations (which also emerged with the understanding of general relativity). Nonetheless, all the mathematical ingredients and physical interpretation were arguably available in the context of Maxwell electromagnetism. Yet the derivation we propose is straightforward only once modern tools and modern understanding of covariant methods are used (at the technical level, the derivation we present uses to a large extend the mathematical tools of general relativity: covariant derivatives and general coordinate invariance). We hope that this paper will present the students with an alternative (perhaps technically more advanced) pedagogical perspective of both technical interest and conceptual value.

The paper is organized as follows. In Section II we give some motivation for our approach by using the heuristics of a photon trapped inside an accelerating box (or a box on a constant gravitational field from the perspective of the equivalence principle). In Section III we review the properties of Rindler coordinates which represent accelerating frames. In Section IV we derive the formula $E = mc^2$ by analysing the energy content of an accelerating box trapping stationary electromagnetic radiation. In Section V we do the same but for a massless scalar field which suggests the universality of our derivation.

In the Appendix we give supplementary material that answers some questions that naturally arise from the our analysis. The only important piece of information for the proof of the main result of the paper, in Section IV, is that the electric field in the rest frame of the box must be perpendicular to the boundary walls (which is obvious for an inertial box but requires justification for an accelerating one). In Appendix A we briefly recall the structure of the covariant version of Maxwells equations and introduce its stress-energy-momentum tensor, we also calculate the properties of electromagnetic field at the boundary of an accelerating box of perfectly conducting walls. In Appendix B, we show that the mass-energy equivalence formula can also be derived from the work done by an external agent accelerating the massless confined radiation. This is the analog of the heuristic argument using the photon given at the beginning of the paper. We do the same for the scalar field in Appendix C.

II. HEURISTICS WITH A TRAPPED PHOTON

As a warming up exercise let us first illustrate the basic idea by using the heuristics provided by the particle interpretation of electromagnetic radiation arising from quantum mechanics. Thus, consider a single photon of frequency ω trapped inside a cubic box of side L. Assume that the box is accelerated with acceleration $|a|$ in the upward direction (see Figure 1). At time $t = 0$ we also assume the box is at rest, and the photon is passing through the center of the box and moving up at the speed of light c. At time

$$\begin{aligned} t_1 &= \frac{L + |a| t_1^2}{2c} = \frac{L}{2c}\left(1 + \mathscr{O}\left(\frac{|a|L}{c^2}\right)\right) \\ &\approx \frac{L}{2c}, \end{aligned} \tag{2}$$

[1] One could speculate that fundamental massive fermions like the electrons might be seen as confined massless radiation as well. Solutions of the Dirac equation can be interpreted as two massless Weyl fermions (the left handed and right handed components of the Dirac fermion) which due to the mutual interaction mediated by the mass term in the Dirac equation constantly annihilate into each other. In this process the momentum of each individual massless component bounces by changing the sign of the momentum in the direction of the spin (Schroedinger's *zitterbewegung* [5]) as if confined in a box of a size of the order of the Compton wave length of the fermion.

the photon hits the top of the box. We are assuming that the speed of the box when the photon hits the top is much smaller that the speed of light, hence $\frac{|a|L}{c^2} \ll 1$. This can happen in two different limits: either the box is very small or the acceleration is very small with respect to the box size.

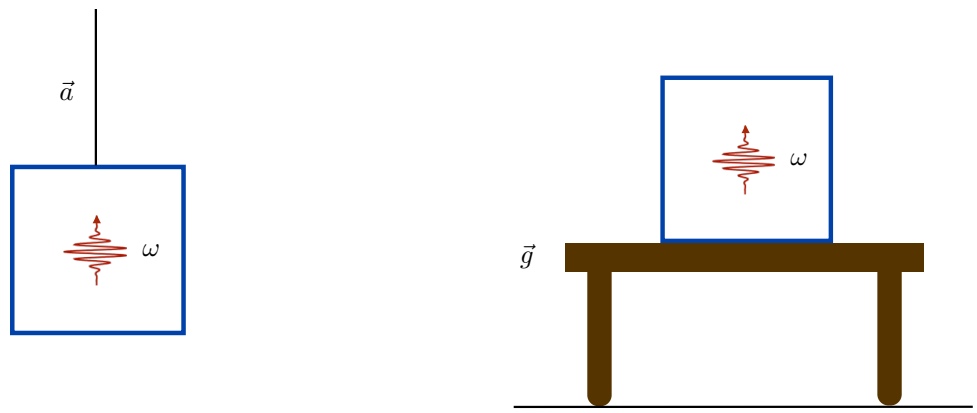

Figure 1: On the left panel: an accelerating box containing a photon. On average, a force $|F| = c^{-2}\hbar\omega|a|$ needs to be applied to maintain the acceleration. The inertia of the box is due to the average of the differential radiation pressure of the photon bouncing on the walls of the box (the pressure is smaller at the top than at the bottom due to the Doppler shift). On the right: the equivalent gravitational situation with the box on a table; the weight of the box is $W = c^{-2}\hbar\omega|g|$. Doppler shift is now replaced by the equivalent gravitational red-shift.

When the photon hits the top its frequency in the rest frame of the top wall is $\omega_1 = (1-|a|t_1/c)\omega$ due to the Doppler shift. Notice that this effect is expected from the non relativistic point of view also: in the limit where $\frac{|a|L}{c^2} \ll 1$ holds, the standard sound-wave-type Doppler red shift formula coincides with the (physically) correct relativistic one. Thus no explicit use of special relativity is being made here. Indeed, in the regime we are working, we can safely assume that time t is always given by the very same inertial time (thought to be absolute time in pre-relativistic terms). At time t_1 the momentum of the photon changes from $\hbar\omega_1/c \to -\hbar\omega_1/c$ in the vertical axis. We have that $\Delta p_1 = -2\hbar\omega_1/c$.

Now, the photon travels downwards and its frequency is back to ω (in the instantaneous rest frame of the box) thanks to the Doppler effect when it passes through the center. Indeed the frequency at the center is always ω as the photons energy must be stationary in the rest frame of the box. This is the easiest understood intuitively by considering the equivalent situation of the box on the table on the right panel of Figure 1. This implies that at time t_2 the photon hits the bottom of the box with a local frequency $\omega_2 = (1 + |a|L/(2c^2))\omega$ and $\Delta p_2 = 2\hbar\omega_2/c$. When the photon gets back to the center at time $t_3 = 2L/c$ the average force $F = \Delta p/\Delta t$ is

$$\begin{aligned} |F| &= \frac{\Delta p}{\Delta t} = \frac{2\hbar\omega}{c}\frac{\left(1 + \frac{|a|L}{2c^2} - 1 + \frac{|a|L}{2c^2}\right)}{2\frac{L}{c}} \\ &= \frac{\hbar\omega}{c^2}|a|, \end{aligned} \tag{3}$$

which implies that the box carries a mass $m = E/c^2$ with $E = \hbar\omega$ (the quantum energy of the photon). The fluctuating character of the mass due to the bouncing back and forth of the photon would go away if we consider many photons in a suitable configuration that makes the radiation inside the box 'stationary' in a way that will become precise in the following section. In the previous heuristic derivation, special relativity is of course hidden in the assumption that the momentum of a photon is $\hbar\omega/c$ but this is really only through quantum mechanics and the dynamical equations of Maxwell theory. That is indeed our point: the mass-energy equivalence relation is encoded in Maxwells dynamics.

The calculation can be improved by considering multiple photons in arbitrary configurations. The final answer remains the same although details become more and more cumbersome. The single photon example suffices as a motivation for what follows. The limitation of the present approach is the use of photons and quantum mechanics. Nevertheless, this simple argument also shows the heart of the reason for the energy-mass equivalence just derived; it resides in the structure of the photon momentum energy relation $p = \hbar\omega/c$ and $E = \hbar\omega$ which is also present in the classical theory, as realized by Einstein, and as pointed out in text books like the classic [6]. More precisely, electromagnetic energy density $\rho \equiv (E^2 + B^2)/(8\pi)$ and the electromagnetic Poynting vector—momentum density

of the electromagnetic radiation—$P_i \equiv (\vec{E}\times\vec{B})_i/(4\pi c)$ are such that for radiation (where $|E| = |B|$ and $\vec{B}\perp\vec{E}$) one has that $|\vec{P}| = E^2/(4\pi c)$ and $\rho = E^2/(4\pi)$ which confirms the energy-mass equivalence via the relationship $(energy/c^2)\times velocity = momentum$ for light. In what follows we will realize this in a clear-cut fashion by making appeal only to the structure of Maxwell equations without needing the details of a particular solution. As long as the radiation is confined inside the box in a stationary configuration the energy-mass formula will follow.

III. RINDLER COORDINATES

We first need to get familiar with the description of an accelerated frame that will be used to represent those observers that are at rest with respect to the idealized box model of a composite particle made of confined electromagnetic radiation (or massless scalar field radiation). We will consider a box full of radiation (electromagnetic fields or massless scalar fields) in Minkowski spacetime whose metric in inertial coordinates takes the standard form

$$ds^2 = -c^2dt^2 + dx^2 + dy^2 + dz^2. \tag{4}$$

Inertial time translations define an isometry of flat spacetime so that $\xi^{\rm lab}\equiv\partial_{ct}$ satisfies a covariant equation known as the Killing equation [2]

$$\nabla_{(a}\xi^{\rm lab}_{b)} = 0. \tag{5}$$

In order to describe the uniformly accelerating box in the x-direction it will be convenient to introduce Rindler coordinates [8] that are related to the inertial coordinates by

$$\begin{aligned} ct &= \bar{x}\sinh(\tau) \\ x &= \bar{x}\cosh(\tau) \end{aligned} \tag{6}$$

so that the flat metric becomes

$$ds^2 = -\bar{x}^2d\tau^2 + d\bar{x}^2 + dy^2 + dz^2. \tag{7}$$

These new coordinates are those associated with uniformly accelerated observers [7]. The inverse transformation is

$$\bar{x} = \sqrt{x^2 - c^2t^2} \tag{8}$$

$$\tau = \operatorname{arctanh}\left(\frac{ct}{x}\right). \tag{9}$$

For later use it is important to write $\xi^a_{\rm lab} = \partial^a_{ct}$ in terms of Rindler coordinates, from (8) we get

$$\begin{aligned} \xi^a_{\rm lab} &= \partial^a_{ct} = \frac{1}{c}\frac{\partial\tau}{\partial t}\partial^a_\tau + \frac{1}{c}\frac{\partial\bar{x}}{\partial t}\partial^a_{\bar{x}} \\ &= \gamma(\tau)\left(\frac{1}{\bar{x}}\partial^a_\tau - \frac{v(\tau)}{c}\partial^a_{\bar{x}}\right), \end{aligned} \tag{10}$$

where we introduced the relativistic gamma factor $\gamma(\tau) = \cosh(\tau) = (1-\beta^2)^{-1/2}$ and $\beta(\tau) = \tanh(\tau) = v(\tau)/c$. Also for later use, the 4-volume form (see for instance Appendix B in [7]) in terms of Rindler coordinates is

$$dv^{(4)} = \bar{x}\, d\bar{x}dydzd\tau, \tag{11}$$

and the 3-volume density for the simultaneity surfaces τ =constant is

$$d\Sigma_\tau = d\bar{x}dydz. \tag{12}$$

Since the metric does not depend on τ,

$$\xi^a_{\rm box} \equiv \partial^a_\tau = \bar{x}\cosh\tau\partial^a_{ct} + \bar{x}\sinh\tau\partial^a_x \tag{13}$$

---

[2] The Killing equation follows from the fact that the Lie derivative of the metric along a vector field defining an isometry vanishes $\mathscr{L}_{\xi^{\rm lab}}\eta_{ab} = 2\nabla_{(a}\xi^{\rm lab}_{b)} = 0$ [7].

is a Killing vector too

$$\nabla_{(a}\xi^{\rm box}_{b)} = 0. \tag{14}$$

The subindex "box" is natural due to the fact that this Killing field is associated with the time translation invariance of the uniformly accelerating observers at rest with the box that will contain the confined radiation as described in the following section (the isometry corresponding to this Killing field is the one associated with the invariance of the flat Minkowski metric under boosts).

The four velocity $u^a_{\rm box}$ of these observers is just proportional the Killing vector, i.e., given by $u^a_{\rm box} = \xi^a_{\rm box}/|\xi_{\rm box}|$, explicitly

$$\begin{aligned} u^a_{\rm box} &= \cosh\tau\partial^a_{ct} + \sinh\tau\partial^a_x \\ &= \gamma(\tau)\partial^a_{ct} + \gamma(\tau)\beta(\tau)\partial^a_x. \end{aligned} \tag{15}$$

The previous expression of $u^a_{\rm box}$ tells us that the stationary observers following the killing trajectories of (13) correspond to orbits of boosts in the x-direction with rapidity given by τ. One can easily compute their acceleration and find that it is constant (independent of τ) and given by

$$\begin{aligned} a^{\rm box}_a &= c^2 u^b_{\rm box}\nabla_b u^{\rm box}_a = c^2\frac{\xi^b_{\rm box}}{|\xi_{\rm box}|}\nabla_b\left(\frac{\xi^{\rm box}_a}{|\xi_{\rm box}|}\right) \\ &= c^2\nabla_a\log(|\xi_{\rm box}|) = c^2\frac{d\bar{x}_a}{\bar{x}}, \end{aligned} \tag{16}$$

where to get the final line we have used the Killing equation (14). Therefore, (15) defines the four velocity field of a box where its bulk points move along constant acceleration trajectories with

$$|a_{\rm box}| = \frac{c^2}{\bar{x}} \tag{17}$$

Notice that even when all points of the box move at the same speed (the box behaves as a rigid box) different points have different acceleration, e.g. the bottom of the box and the top of the box accelerate differently so that the box remains un-stretched (the distance between the top and the bottom of the box remains fixed). This might be surprising at first sight but it is one of these counter intuitive facts in Lorentzian geometry. All the same, when the concept of acceleration of the box will be needed (only in the material presented in the Appendices) we will work under the assumption that $\frac{|a_{\rm box}|L}{c^2} \ll 1$ (already used in Section II) in which case a single constant notion of acceleration can be assigned in an approximate manner to the whole box.

IV. THE MASS OF MAXWELL FIELDS CONFINED IN A BOX

In this section we derive the mass-energy equivalence formula from the inertial properties of a box made of perfectly conducting walls filled with electromagnetic radiation. We consider Maxwell theory and its solutions in Rindler coordinates, introduced in the previous section, and impose the boundary conditions that represent the presence of perfectly conducting walls in that frame. However, we will see that very little information about the solutions is needed (and this is one of the nice features of the result). More precisely, the only explicit thing that enters the proof below is that the electric field, in the rest frame of the accelerating box, must be perpendicular to the walls of the box. This is well known for a box in inertial motion, the fact that it remains true in the uniformly accelerated case is perhaps physically clear but technically less obvious. The proof is given in the Appendix A 1 whose main results are discussed in A 2, and the basic mathematical reason is that Maxwell's equations maintain very much the same structure on the accelerated frame as in an inertial one.

Let us now compute the energy content of the box, as measured in the laboratory frame, at a given simultaneity surface of constant Rindler time τ (see Figure 2). In order to do this we introduce the stress-energy-momentum tensor (energy-momentum tensor from now on) T_{ab} of the Maxwell field and the current

$$j^{\rm lab}_a \equiv -T_{ab}\xi^b_{\rm lab}. \tag{18}$$

The energy momentum tensor for electromagnetism is given in (A5); however, its explicit form is not important at the moment. By definition of the current in the previous equation, the energy content of the box of confined radiation at a given τ as measured in the lab frame is

$$E_{\rm box}(\tau) = -c^2\int_{\Sigma_\tau} j^{\rm lab}_a n^a\, d\Sigma_\tau, \tag{19}$$

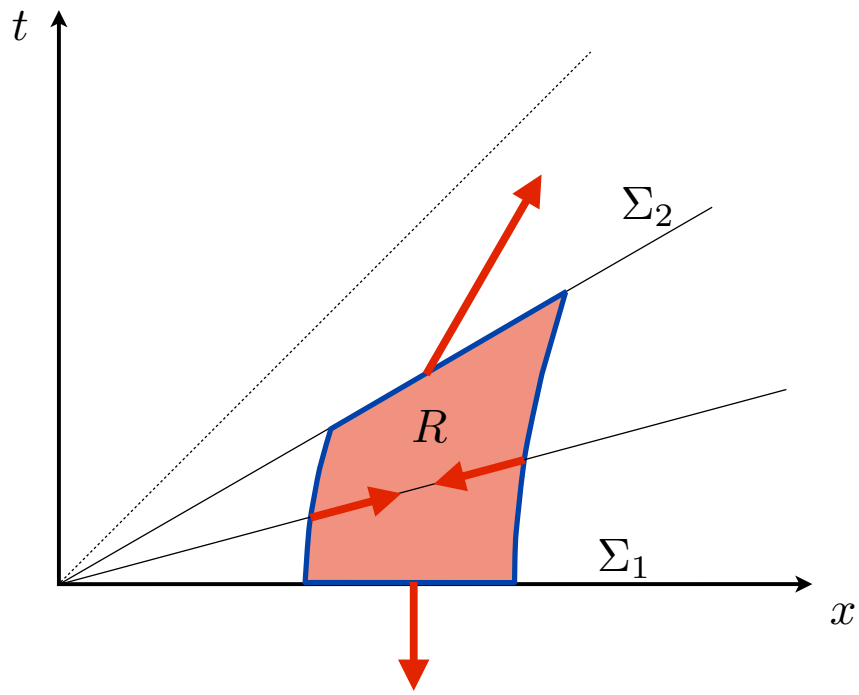

Figure 2: Accelerated box in Minkowski space-time. Σ represent constant time surfaces while R is the region inside the box. The hyperbola correspond to the trajectory of the walls perpendicular to the motion.

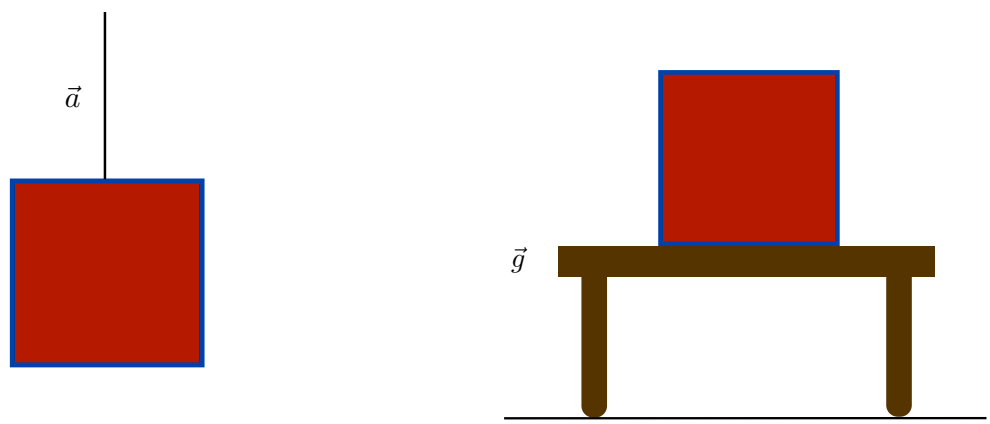

Figure 3: A box of accelerated radiation on the left. On the right the equivalent situation of radiation in a box near the vicinity of the earth. The energy-mass equivalence holds for the confined radiation. The only assumption is the stationarity of the radiation in the rest frame of the box and the validity of the field equations.

where $d\Sigma_\tau = d\bar{x}dydz$ is the volume density of the hypersurface τ =constant as introduced in (12) and n^a is the normal to these hypersurfaces. Explicitly, $n^a = \bar{x}^{-1}\partial^a_\tau$ which when replaced in the previous equation gives

$$E_{\rm box}(\tau) \;=\; -c^2\int_{\Sigma_\tau} j^{\rm lab}_a n^a d\Sigma_\tau = c\int \bar{x}^{-1}T_{t\tau}d\bar{x}dydz. \tag{20}$$

Now, from (10) we get

$$E_{\rm box}(\tau) \;=\; \gamma(v)c^2\int_{\Sigma_\tau}\bar{x}^{-2}T_{\tau\tau}d\bar{x}dydz - vc\gamma(v)\int_{\Sigma_\tau}\bar{x}^{-1}T_{\tau\bar{x}}d\bar{x}dydz. \tag{21}$$

As we show below, the second term in the previous equation vanishes if we demand that the radiation inside the

box is stationary in its rest frame, namely that the space components of the total linear momentum of the radiation vanishes in the rest frame of the box. Strictly speaking we only need the total momentum in the $\bar{x}$ direction to vanish. However, assuming we want our box to represent a simple model of composite particle that could be accelerated in arbitrary directions then the vanishing of all component of the space part of the linear momentum in the rest frame of the box will be a natural demand. In order to show that this is equivalent to the vanishing of the second term in (21), let us now consider the energy current associated to the frame in which the box is at rest. So we have

$$j_a^{\rm box} = -T_{ab}u^b_{\rm box} = -\bar{x}^{-1}T_{ab}\partial^b_\tau, \tag{22}$$

where $u^a_{\rm box} = \xi^a_{\rm box}/|\xi_{\rm box}| = \bar{x}^{-1}\partial_\tau$ as given in (15) is the four velocity of the box. The momentum density of the radiation inside the box along the direction $\partial_{\bar{x}}$ as measured by the observer $u^a_{\rm box}$ at a time τ is given by $p_{\hat{x}}(\tau) = j_a^{\rm box}\partial^a_{\bar{x}}$. The condition for the vanishing of the total momentum of the radiation in the rest frame of the box in the $\bar{x}$ direction takes the form [3]

$$\int_{\Sigma_\tau} j_a^{\rm box}\partial^a_{\bar{x}}dv^{(3)} = -\int_{\Sigma_\tau}\bar{x}^{-1}T_{\tau\bar{x}}d\bar{x}dydz = 0. \tag{23}$$

The previous is the only real requirement on the solutions of the electromagnetic field inside of the confining box. It has a natural physical meaning corresponding to restricting the radiation to a stationary configuration which reflects the notion of a compact particle-like object that we have in mind. In more general situations the two terms in (21) are important.

The previous stationarity condition reduces (21) to

$$E_{\rm box}(\tau) \;=\; \gamma(\tau)\int_{\Sigma_\tau}\frac{c^2}{\bar{x}^2}T_{\tau\tau}d\bar{x}dydz \tag{24}$$

Now we show that the integral in the previous equation does not depend on Σ_τ. A simple calculation of the divergence of the current (22) yields

$$\begin{aligned}\nabla^a j_a^{\rm box} = -\nabla^a\left(\bar{x}^{-1}T_{ab}\xi^b_{\rm box}\right) &= \bar{x}^{-1}\sigma^a E_a^{\rm box}\delta_{\rm box} - T_{a\tau}g^{ac}\nabla_c\bar{x}^{-1}\\ &= \bar{x}^{-2}T_{\tau\bar{x}}\end{aligned} \tag{25}$$

where in the first line we have used (14) and (A8), and in the second line the fact that $\sigma^a E_a^{\rm box} = 0$ (the electric field in the box frame must be orthogonal to the surface current for perfectly conducting walls). Indeed, the electric field $E_a^{\rm box}$ at the walls of the box is orthogonal to the walls of the box while the normal component of the magnetic field $B_a^{\rm box}$ vanishes at the walls. Even though this might be physically clear, the mathematical proof from Maxwell equations is tricky because one is on a non-inertial frame. We present it in the Appendix A 1 and A 2.

As implied by (25) the current $j_a^{\rm box}$ is not locally conserved; nevertheless, when we integrate it in space-time region R swept by the box (see figure 2) we find

$$\int_R \nabla^a j_a^{\rm box}dv^{(4)} \;=\; \int d\tau\left(\int_{\Sigma_\tau}\bar{x}^{-1}T_{\tau\bar{x}}d\bar{x}dydz\right) = 0, \tag{26}$$

where we used (11) and the last integration vanishes because the quantity in the parenthesis vanishes due to the stationarity condition (23). Now, it follows from the perfect conductor boundary conditions that

$$j^a_{\rm box}N_a|_{\rm walls} = 0, \tag{27}$$

where N^a is the normal to the walls of the box (see detail proof in Appendix A 2, equation (A29)). As a result, Gauss theorem implies that the flux across the boundary of the region R receives only contributions from the spacelike components of the boundary of R (see in Figure 2), namely

$$0 = \int_R \nabla^a j_a^{\rm box}dv^{(4)} = \int_{\Sigma_2} j_a^{\rm box}u^a_{\rm box}d\Sigma_2 - \int_{\Sigma_1} j_a^{\rm box}u^a_{\rm box}d\Sigma_1, \tag{28}$$

[3] Notice that as the previous integral involves the notion of '$\bar{x}$-direction' at different point in the box, one would need in principle to parallel transport all these vectors at some reference point to be able to integrate (sum up) all the contributions. It is easy to check that such parallel transport is trivial in this case.

hence the integration of $j_a^{\rm box} u^a_{\rm box}$ does not depend on the τ =constant hypersurface Σ_τ: it is a constant of motion. What is the physical interpretation of that constant? From the fact that $u^a_{\rm lab} = u^a_{\rm box}$ at $\tau = 0$ we see that this constant is nothing else but the rest energy of the radiation in the box

$$E(0) = -\int_\Sigma c^2 j_a^{\rm box} u^a_{\rm box} d\Sigma = \int_\Sigma c^2 T_{ab} u^a_{\rm box} u^b_{\rm box} d\Sigma = \int_\Sigma c^2 \bar{x}^{-2} T_{\tau\tau} d\bar{x} dy dz. \tag{29}$$

Therefore, equation (24) takes the form

$$E_{\rm box}(v) = \gamma(v) E_{\rm box}(0), \tag{30}$$

where we are now using the direct correspondence $\tau = \operatorname{arctanh}(v/c)$ and hence trading τ by the velocity v in the previous expression. Expanding $\gamma(v)$ to leading order on v we find

$$E_{\rm box}(v) = \left(1 + \frac{v^2}{2c^2}\right) E_{\rm box}(0) + \mathscr{O}\left(\frac{v^4}{c^4}\right) E_{\rm box}(0). \tag{31}$$

Correspondence with the non relativistic limit requires

$$\boxed{E_{\rm box}(0) = mc^2} \tag{32}$$

where m is the rest mass of the confined electromagnetic radiation.

V. THE MASSLESS SCALAR FIELD CASE

The result of the previous section depends of the field equations only and in a generic manner, in the sense that no particular solutions need to be considered for the proof. The first important ingredient is the conservation of the energy-momentum tensor in the bulk of the box (which follows from the validity of the field equations). The second is the behaviour of the divergence of $j_a^{\rm box}$ in (25) where the reflecting boundary conditions constrain the electric field (as measured in the box frame) to be orthogonal to the walls, and the third is the orthogonality of $j_a^{\rm box}$ to the walls. Both these ingredients follow from the structure of the field equations at the boundary as shown in Appendix A 2. But the important things is that no specific solution needs to be chosen to prove these properties: these are generic consequences of the equations and the physical conditions at the idealized walls of the box. Therefore, one would expect that the proof of the previous section should be valid (with small adjustments) for any massless matter model. We do not have a general proof of this; nevertheless, we can at least build up evidence by exhibiting another simple example: the massless scalar field.

The field equation of a massless scalar field ϕ is

$$\Box\phi \equiv g^{ab}\nabla_a\nabla_b\phi = j, \tag{33}$$

where j is a source term (necessary to impose the boundary conditions at the walls of the box). The stress-energy-momentum tensor is given by

$$T_{ab} = \nabla_a\phi\nabla_b\phi - \frac{1}{2} g_{ab}\ g^{cd}\nabla_c\phi\nabla_d\phi. \tag{34}$$

Direct calculation of the divergence of the stress-energy-momentum tensor (34) yields

$$\nabla^a T_{ab} = j\nabla_b\phi, \tag{35}$$

which in the absence of sources vanishes identically. When the radiation is confined inside a box, made of perfectly reflecting walls (as in the Maxwell case) surface charges appear. We write

$$j = \sigma\delta_{\rm box}, \tag{36}$$

where σ represent the surface charge density. This is the analog of the surface electric charges and surface current in a perfect conducting wall in electromagnetism. They are the responsible of imposing reflecting boundary conditions that, in the present case, boil down to $\phi = 0$ at the box walls. These sources fix the normal derivative of the scalar field: from (33), and the Gauss law applied to the vector field $\nabla^a\phi$, it follows that

$$N^a\nabla_a\phi = \sigma. \tag{37}$$

or simply

$$\nabla_a \phi = \sigma N_a. \tag{38}$$

As in the case of Maxwell fields we start form the definition of the energy content of the box in the lab frame (19). The argument follows the same lines from (19) to (24) where the field equations do not really enter. Things change slightly when considering the current (22) whose divergence remains

$$\begin{aligned} \nabla^a j_a^{\rm box} = -\nabla^a \left(\bar{x}^{-1} T_{ab} \xi^b_{\rm box}\right) &= -\bar{x}^{-1}\sigma(u^a_{\rm box}\nabla_a\phi)\delta_{\rm box} - T_{a\tau} g^{ac}\nabla_c \bar{x}^{-1} \\ &= \bar{x}^{-2} T_{\tau\bar{x}}, \end{aligned} \tag{39}$$

due to the fact that $u^a_{\rm box}\nabla_a\phi = \bar{x}^{-1}\partial_\tau\phi = 0$ at the boundary (either because $\phi = 0$ for all τ or, equivalently, due to equation (38) and the fact that $u^a_{\rm box}N_a = 0$). Therefore, the equivalent of equation (26) is also valid for the scalar field as long as the stationarity condition (23) is satisfied for the scalar field inside the box. Now, the validity of equation (28) depends on the validity of $j_a^{\rm box}N^a = 0$. Using the definition of the energy-momentum tensor and the fact that $u^a_{\rm box}N_a = 0$ we see that

$$j_a^{\rm box} N^a\big|_{\rm walls} = -(u^a_{\rm box}\nabla_a\phi)(N^b\nabla_b\phi) = 0, \tag{40}$$

due to the boundary condition $u^a_{\rm box}\nabla_a\phi = \bar{x}^{-1}\partial_\tau\phi = 0$. The rest of the argument from (28) to the main result (32) now follows exactly as in the Maxwell case.

VI. CONCLUSIONS

We have shown that confined radiation in an idealized box with walls imposing perfectly reflecting boundary conditions for both Maxwell electromagnetic fields and massless scalar fields has an inertial mass given by its energy content divided by the square of the speed of light. Our calculation relies entirely on general properties of the solution of the field equations and the properties of the energy momentum tensor of the confined radiation. The only explicit requirement on the solutions is that the radiation be in a stationary state of vanishing total linear momentum in the frame of the box (not moving inside the box). This assumption is compatible with the idea of the box representing a toy-model of a composite particle (an ultra simplified classical model of proton or a neutron). The calculations done in this solvable simple model of a composite particle has deep conceptual implications making natural the possibility that all mass parameters in our physical models could have a more fundamental description in terms of more basic degrees of freedom (mass as an emergent notion).

The proof of main claim is straightforward once the relevant equations are written in covariant form. Even when no gravitational field is invoked, the result follows naturally from the application of the mathematics of general relativity. On the physical front, the naturalness of the energy-momentum tensor as the source of gravity is made more transparent by our pre-gravitational analysis. For those reason we expect the paper to be useful from a pedagogical perspective.

VII. ACKNOWLEDGEMENTS

The basic idea of this paper started in discussion with a group students of the first year master in theoretical physics at Aix-Marseille University. We thank discussions with F. Balfour, M. L. Frisch Sbarra, A. Vesperini, and S. Charfi.

Appendix A: Maxwell equations and boundary conditions

Maxwell equations in covariant form and in the presence of sources are

$$\nabla^a F_{ab} = -4\pi J_b \tag{A1}$$

$$\nabla_a F_{bc} + \nabla_b F_{ca} + \nabla_c F_{ab} = 0. \tag{A2}$$

where $F_{ab} = -F_{ba}$ is the electromagnetic field strength, and J_a is the electric four-current. For an arbitrary observer with four velocity u^a the electric field is given by

$$E_a = F_{ab}u^b \tag{A3}$$

while the magnetic field is

$$B_a = -\frac{1}{2}\epsilon_{abcd}F^{cd}u^b. \tag{A4}$$

The stress-energy-momentum tensor of the electromagnetic field is

$$T_{ab} = \frac{1}{4\pi}(F_{ac}F_b{}^c - \frac{1}{4}g_{ab}\ F_{cd}F^{cd}). \tag{A5}$$

It follows from the validity of (A1) and (A2) that the divergence of T_{ab} is given by

$$\nabla^a T_{ab} = J^a F_{ab}. \tag{A6}$$

We will assume that the electromagnetic field is confined in a box without charges inside. Therefore, $J^a = 0$ in the bulk of the box. However, boundary currents must be present to ensure that the fields vanish outside the confining box (they are responsible for enforcing reflecting boundary conditions). We assume that the walls are made of a perfect conductor with infinitely light charge carriers that can move freely. We will hence write the current as

$$J^a = \sigma^a \delta_{\text{box}}, \tag{A7}$$

where σ^a is the surface current and δ_{box} denotes the Dirac distribution with support on the walls of the box. Thus from (A6) we have

$$\nabla^a T_{ab} = \sigma^a F_{ab}\delta_{\text{box}}. \tag{A8}$$

The energy-momentum current associated to the lab-frame ($\xi^a_{\text{lab}} = \partial^a_{ct}$) is

$$j^{\text{lab}}_a \equiv -T_{ab}\xi^b_{\text{lab}} \tag{A9}$$

which is not conserved because of the contributions of the boundary degrees of freedom mentioned above. In fact, from Maxwell equations we get

$$\nabla_a j^a_{\text{lab}} = -J^a F_{ab}\xi^b_{\text{lab}} = -\sigma^a E^{\text{lab}}_a \delta_{\text{box}} \tag{A10}$$

where we used that $\nabla_a \xi^{\text{lab}}_b + \nabla_b \xi^{\text{lab}}_a = 0$ because ξ^{lab}_a is a Killing field, recall (5). Note that the right hand side of the previous equation vanishes inside the box where $J^a = 0$. If the box is at rest then $\sigma^a E^{\text{lab}}_a = 0$ on the boundary due to the perfect conductor boundary conditions[4] and the current is conserved. However, one can have $\sigma^a E^{\text{lab}}_a \neq 0$ in general situations where the box is moving; such possibility is important and plays a role in Section B.

1. Maxwell equations in the accelerated frame of the box

A specially interesting case for the present paper is the one corresponding to a uniformly accelerated box. Thus we analyse the content of Maxwell theory in terms of the electric and magnetic fields defined in an accelerating frame. Given the four velocity of a family of observers at rest with respect to the accelerating box u^a_{box}–recall (15)–one can write the electromagnetic field strength as

$$F_{ab} = -2E^{\text{box}}_{[a}u^{\text{box}}_{b]} - \epsilon_{abcd}B^c_{\text{box}}u^d_{\text{box}} \tag{A11}$$

where $E^{\text{box}}_a = F_{ab}u^b_{\text{box}}$ and $B^{\text{box}}_a = -\frac{1}{2}\epsilon_{abcd}F^{cb}u^b_{\text{box}}$ [5]. It follows from the skew symmetry of F_{ab} that

$$B^{\text{box}}_a u^a_{\text{box}} = 0 = E^{\text{box}}_a u^a_{\text{box}}. \tag{A12}$$

Now we write Maxwell equations (A1) in a way that it would lead to the analog of Gauss and Ampére integral identities for inertial frames but now these are valid in an accelerated frame. This step is rather technical but very important; a general treatment in curved spacetimes can be found in [9].

---

[4] Otherwise the charges would accommodate as they can move due to the electric force and neutralise any parallel components of the electric field.

[5] Indeed, the previous expression is valid for any timelike vector field u^a of four-velocities representing a field of observers in spacetime.

First notice that $u_a^{\rm box} = g_{ab}u^a_{\rm box} = -\bar{x}\nabla_a\tau$. This suggests the introduction of a new quantity, $\overline{u}_a \equiv -\nabla_a\tau$, which has the following nice properties:

$$\begin{aligned} \nabla_a\overline{u}^a &= 0 \\ \nabla_a\overline{u}_b &= \nabla_b\overline{u}_a \\ h^c_a h^d_b\nabla_c\overline{u}_d &= 0, \end{aligned} \tag{A13}$$

where $h_{ab} = g_{ab} + u^{\rm box}_a u^{\rm box}_b$ is the spacial metric of the box simultaneity slices, τ =constant slices in Figure 2. The first property says that the $\overline{u}^a$ congruence is divergence free, the second implies that it is surface forming (trivially coming from the fact that $\overline{u}_a$ is an exact form normal to the τ =constant surfaces), and the last property implies that it is shear free [7]. Equation (A11) can be written as

$$F_{ab} = -2\overline{E}_{[a}\overline{u}_{b]} - \epsilon_{abcd}\overline{B}^c\overline{u}^d \tag{A14}$$

where $\overline{E}^a = \bar{x}E^a_{\rm box}$ and $\overline{B} = \bar{x}B^a_{\rm box}$. Maxwell equation (A1) becomes

$$\begin{aligned} -4\pi J_b &= -\nabla^a(\overline{E}_a\overline{u}_b) + \nabla^a(\overline{E}_b\overline{u}_a) - \epsilon_{abcd}\nabla^a(\overline{B}^c\overline{u}^d) \\ &= -(\nabla^a\overline{E}_a)\overline{u}_b - \overline{E}^a\nabla_a\overline{u}_b + \overline{u}^a\nabla_a\overline{E}_b - \epsilon_{abcd}(\nabla^a\overline{B}^c)\overline{u}^d \\ &+ \underbrace{\overline{E}_b\nabla^a\overline{u}_a - \epsilon_{abcd}\overline{B}^c(\nabla^a\overline{u}^d)}_{=0}, \end{aligned} \tag{A15}$$

where for the moment we just used the Leibniz rule and wrote at the end the two terms that vanish identically due to the first two identities in (A13). The next step is to separate the previous equation into its part parallel to $u^a_{\rm box}$ (projecting with $-u^a_{\rm box}u^{\rm box}_b$) and its normal or spacial part (which we can obtain by projecting with $h^a_b = \delta^a_b + u^a_{\rm box}u^{\rm box}_b$).

Before doing the projections we notice that

$$\begin{aligned} -4\pi J_b &= -(\nabla^a\overline{E}_a)\overline{u}_b - \overline{E}^a\nabla_a\overline{u}_b + \overline{u}^a\nabla_a\overline{E}_b - \overbrace{\epsilon_{abcd}(\nabla^a\overline{B}^c)\overline{u}^d}^{\text{orthogonal to } \overline{u}^a} \\ &= -(\nabla^a\overline{E}_a)\overline{u}_b - \overline{E}^a\nabla_b\overline{u}_a + \overline{u}^a\nabla_a\overline{E}_b - \epsilon_{abcd}(\nabla^a\overline{B}^c)\overline{u}^d \\ &= -(\nabla^a\overline{E}_a)\overline{u}_b + \overline{u}^a\nabla_b\overline{E}_a + \overline{u}^a\nabla_a\overline{E}_b - \epsilon_{abcd}(\nabla^a\overline{B}^c)\overline{u}^d \end{aligned} \tag{A16}$$

where in the second line we used the second equation in (A13) for the second term, and in the third line we used that $\overline{E}_a\overline{u}^a = 0$, or (A12). Let us now project along $u^a_{\rm box}$ recalling that $u^a_{\rm box}\overline{u}_a = u^a_{\rm box}(u^{\rm box}_a\bar{x}^{-1}) = -\bar{x}^{-1}$ we get

$$\begin{aligned} -4\pi\bar{x}J_b u^b_{\rm box} &= (\nabla^a\overline{E}_a) + 2u^a_{\rm box}u^b_{\rm box}\nabla_b\overline{E}_a \\ &= (g^{ab} + u^a_{\rm box}u^b_{\rm box})\nabla_b\overline{E}_a + u^a_{\rm box}u^b_{\rm box}\nabla_b\overline{E}_a \\ &= \underbrace{(g^{ab} + u^a_{\rm box}u^b_{\rm box})\nabla_b\overline{E}_a}_{\equiv D^a\overline{E}_a} - (u^b_{\rm box}\nabla_b u^{\rm box}_a)\overline{E}^a, \end{aligned} \tag{A17}$$

where in the last line we used $\overline{E}_a\overline{u}^a = 0$ again and we have used the definition of the spacial covariant derivative D_a such that $D_a h_{bc} = 0$ [7]. Substituting the expression (16) of the acceleration, and $\overline{E}_a = \bar{x}E^{\rm box}_a$ in the last equation we obtain the familliar Gauss law

$$-4\pi J_b u^b_{\rm box} = \frac{1}{\bar{x}}D^a(\bar{x}E^{\rm box}_a) - \frac{D_a\bar{x}}{\bar{x}}E^a_{\rm box} \tag{A18}$$

simplifying

$$\boxed{-4\pi J_a u^a_{\rm box} = D^a(E^{\rm box}_a),} \tag{A19}$$

which has the form of the usual Gauss law in an inertial frame. Indeed it is easy to show that the Gauss law holds in its usual form in arbitrary frames (see Problem 2 in Chapter 4 of [7]). In the present case the technical complications of the previous lines are justified not by the objective of obtaining the Gauss law but rather the aim of getting the analog of Ampére's law which will follow from the spacelike part of the previous equations.

Therefore, we need to project (A16) using $h^a_b = \delta^b_c + u^b_{\rm box}u^{\rm box}_c$. But before we notice that the first term projects to zero while the last term projects to itself. Let us analyse the remaining terms before projecting. There is

$$\overline{u}^a\nabla_b\overline{E}_a = -\overline{E}^a\nabla_b\overline{u}_a, \tag{A20}$$

which in its form on the right clearly projects to zero according to the third equation in (A13) and the fact that $\overline{E}^a$ is purely spacelike. Now let us analyse the remaining term

$$\overline{u}^a\nabla_a\overline{E}_b = \bar{x}^{-1}(\underbrace{u^a_{\rm box}\nabla_a\overline{E}_b + \overline{E}_a\nabla_b u^a_{\rm box}}_{\mathscr{L}_{u_{\rm box}}\overline{E}_b} - \overline{E}_a\nabla_b u^a_{\rm box}), \tag{A21}$$

where we have added and subtracted the same term on the right just to recover the expression of the Lie derivative $\mathscr{L}_{u_{\rm box}}\overline{E}_b$ which is the natural derivative along the world-lines of the box observers. Notice that the term we added and subtracted projects to zero (purely time-like) due to the third equation in (A13). The Lie derivative in the previous equation corresponds to a natural proper time $\mathrm{T} \equiv \bar{x}\tau$ derivative of the electric $\overline{E}_a$. Its space projection is the proper time Fermi transport derivative [9], we denote this

$$D_{\rm T}\overline{E}_a \equiv h^b_a(\mathscr{L}_{u_{\rm box}}\overline{E}_b) = h^b_a(\overline{u}^c\nabla_c\overline{E}_b), \tag{A22}$$

where the previous equivalence of derivatives is valid in our simple case due to (A13). For the general relationship among these see [9]. Thus, finally putting all this together and projecting into the space part of (A16) we get

$$\boxed{-4\pi\bar{x}J^{\rm space-part}_b = D_{\rm T}(\bar{x}E^{\rm box})_b - (D\times\bar{x}B^{\rm box})_b}, \tag{A23}$$

where (as in A19) D_a is the 3d covariant derivative compatible with the space metric h_{ab}. Finally, the homogeneous Maxwell equations (A2) can be written as

$$\nabla^a F^\star_{ab} = 0 \tag{A24}$$

where

$$F^\star_{ab} = \frac{1}{2}\epsilon_{abcd}F^{cd} = 2B^{\rm box}_{[a}u^{\rm box}_{b]} + \epsilon_{abcd}E^c_{\rm box}u^d_{\rm box}. \tag{A25}$$

The previous is the analog of F_{ab} as given in (A11) where $B_a \to -E_a$. As $B_a u^a_{\rm box} = 0$ as well and this was the only requirement entering the derivation of (A23) and (A19) in addition to the properties of $\overline{u}_a$ (A13), it follows from (A24) that

$$\boxed{D^a(B^{\rm box}_a) = 0,} \tag{A26}$$

and

$$\boxed{-D_{\rm T}(\bar{x}B^{\rm box})_b + (D\times\bar{x}E^{\rm box})_b = 0.} \tag{A27}$$

Equations (A23), (A19), (A26), and (A27) are Maxwell's equations for the electric and magnetic field on the accelerated (instantaneous rest) frame of the confining box.

2. Consequences

In the previous section we have recast the Maxwell equations in terms of the electric and magnetic fields as measured in the rest frame of the accelerating box. It was a bit technical but the consequences for the electromagnetic field near the perfectly conducting walls of the box are quite simple and analogous to those that one is familiar with for a box at rest in an inertial frame. In this short section we analyse and state them. We will now see that, as in the case of an inertial box, the Maxwell equations (plus the standard physical assumption that the magnetic field inside the conductor is initially zero) applied to the accelerating box imply that

$$F_{ab}(\text{inside conductor}) = 0. \tag{A28}$$

More precisely, the idealization of perfectly conducting walls requires first the electric field to vanish inside the conductor, $E^{\rm box}_a(\text{inside conductor}) = 0$. In addition, right inside the box and at the walls any parallel component of $E^{\rm box}_a$ to the walls must vanish: if not there would be a force rearranging surface charges to make this component vanish

[6]. Therefore $E_a^{\rm box}|_{\rm box} \propto N_a$ where N_a is the normal to the walls. Now, equation (A27) implies that the magnetic field $B_a^{\rm box}$ must be time independent inside the conductor. Assuming that the magnetic field was zero initially then we have that $B_a^{\rm box}$(inside conductor) $= 0$ for all times. Equation (A28) now follows from (A11).

Another important consequence of the vanishing of the magnetic field inside the box is that from (A26) one can prove that the normal component of $B_a^{\rm box}$ at the walls (on the inside of the box) must vanish. An important consequence of this follows from a two lines calculation that uses (A11), the definition $j_a^{\rm box} \equiv -T_{ab}u^b_{\rm box}$, and (A5), and leads to the important equation

$$j_a^{\rm box} N^a\big|_{\rm walls} = \frac{1}{4\pi}\,\vec{B}_{\rm box}\cdot(\vec{E}_{\rm box}\times\vec{N})\Big|_{\rm walls} = 0, \tag{A29}$$

where we have used once again that $E_a^{\rm box}|_{\rm box} \propto N_a$.

In the two sections that follow we will use (A28) and Maxwell's equations to express F_{ab} at the walls of the box explicitly in terms of the surface charge current. This will then allow us to write explicitly the energy-momentum tensor (A5) at the walls which is important in the analysis of Appendix B. We do this first using three-dimensional methods that involve Gausses law and Ampère's law and later in a more direct covariant fashion.

3. Electromagnetic field at the boundary: canonical derivation

Let us consider the case of a box made of perfectly conducting walls. Then the presence of surface charges is characterized by the four-current

$$J^a = \sigma^a \delta_{\rm box} \tag{A30}$$

As charges can move freely on the walls, only the normal component of the electric field is non vanishing at the wall. One can use this fact and the Gauss law (the integral form of (A19) using a suitably chosen region) at the wall and obtain

$$E_a^{\rm box}\big|_{\rm walls} = -4\pi\,(\sigma_b u^b_{\rm box})\big|_{\rm walls}\,N_a, \tag{A31}$$

where N^a is the unit normal to the wall. Similarly, using Ampère's law (Stokes theorem and (A23)), and the fact that the magnetic field has only parallel components to the walls, one obtains

$$B_a^{\rm box}\big|_{\rm walls} = -\,4\pi\epsilon_{abcd}\sigma^b N^c u^d_{\rm box}\big|_{\rm walls}\,. \tag{A32}$$

In order to prove the previous statement one chooses an infinitesimal 2-surface transversal to the walls and such that its normal is aligned with the surface current. This choice and equation (A23) yields immediately (A32); the time derivative term in (A23) does not contribute because it is orthogonal to the surface's normal.

With this the field-strength (A11) on the walls of the box is given by

$$\boxed{F_{ab} = 8\pi N_{[a}\sigma_{b]}.} \tag{A33}$$

4. The energy momentum tensor at the walls

With the previous result we can now write the energy momentum tensor at the walls of the box using its definition (A5)

$$T_{ab}|_{\rm box} \;=\; 4\pi\left(\sigma_a\sigma_b + (\sigma\cdot\sigma)N_aN_b - \frac{g_{ab}}{2}(\sigma\cdot\sigma)\right), \tag{A34}$$

where we have used the boundary condition $N^a\sigma_a = 0$. It follows that

$$j_a^{\rm lab} N^a\big|_{\rm box} = -2\pi(\sigma\cdot\sigma)N_b\xi^b_{\rm lab}, \tag{A35}$$

---

[6] Here we are assuming idealized charge carriers without mass. Real massive charges would produce a parallel $E_a^{\rm box}$ component to equilibrate for the gravitational pull as it is intuitive from the perspective offered by the right panel in Figure 3.

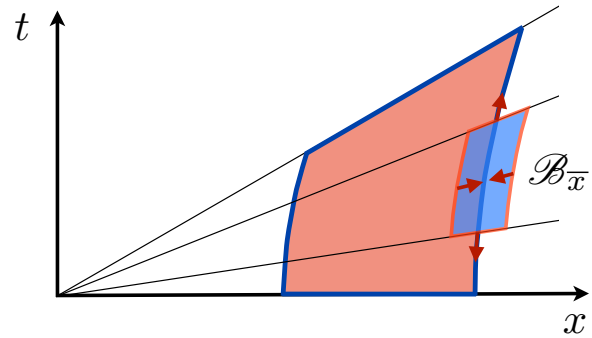

Figure 4: Four-dimensional region $\mathscr{B}_{\bar{x}}$ at the boundary of the world-tube of the box where the application of Gauss theorem leads directly to (A33).

and from (A10) and (A33)

$$\nabla^a j_a^{\rm lab} = -\sigma^a F_{ab}\xi^b_{\rm lab}\big|_{\rm box}\,\delta_{\rm box} = 4\pi(\sigma\cdot\sigma)N_b\xi^b_{\rm lab}\delta_{\rm box}. \tag{A36}$$

Finally, notice that the pressure at the walls is given by

$$P_N \equiv T_{ab}N^aN^b = 2\pi(\sigma\cdot\sigma). \tag{A37}$$

5. Electromagnetic field at the boundary: covariant derivation

Equation (A33) was derived from the Gauss and Ampère's laws on an accelerating box in a way that follows the standard text books type of considerations in inertial frames that break covariance by invoking the electric and magnetic fields. However, the simplicity of (A33) calls for a more direct and covariant derivation. As an exercise, here we show that such more direct path is actually available.

The key input is the requirement that charges on the wall move freely and so cancel any parallel component of the electric field $E_a^{\rm box} = F_{ab}u^b_{\rm box}$ on the rest frame of the box, and similarly that the normal magnetic field component vanishes at the wall (as shown in Appendix A 2). Consider the four vectors $X^a \equiv \nabla^a\bar{x}$, $Y^a \equiv \nabla^a y$, $Z^a \equiv \nabla^a z$, and $T^a \equiv \nabla^a\tau$ which by definition are such that

$$\nabla^{[a}X^{b]} = \nabla^{[a}Y^{b]} = \nabla^{[a}Z^{b]} = \nabla^{[a}T^{b]} = 0. \tag{A38}$$

The previous set of equations together with the definition of the currents $p_a^Y = F_{ab}Y^b$, $p_a^Z = F_{ab}Z^b$, and $p_a^T = F_{ab}T^b$ imply, from (A1), that

$$\nabla^a p_a^Y = -4\pi J_a Y^a, \qquad \nabla^a p_a^Z = -4\pi J_a Z^a, \qquad \nabla^a p_a^T = -4\pi J_a T^a. \tag{A39}$$

Applying the Gauss theorem to p_a^Y in the region $\mathscr{B}_{\bar{x}}$ shown in figure 4 we get

$$\begin{aligned}\int_{\mathscr{B}_{\bar{x}}} \nabla^a p_a^Y &= \int_{\partial\mathscr{B}_{\bar{x}}} p_a^Y n^a \\ -4\pi\int_{\mathscr{B}_{\bar{x}}} J_a Y^a &= \int_{\partial\mathscr{B}_{\bar{x}}} F_{ab}n^aY^b,\end{aligned} \tag{A40}$$

where n^a is the normal to the boundary $\partial\mathscr{B}_{\bar{x}}$ with the orientation shown in Figure 4. Notice that as the normal to the bottom and top (spacelike) portions of the boundaries are proportional to $u^a_{\rm box}$ one has there that $F_{ab}n^aY^b \propto E_a^{\rm box}Y^a = 0$ as only the normal component along X^a of the rest-frame electric field $E_a^{\rm box}$ is non vanishing due to the presence of perfectly conducting walls. In addition $F_{ab} = 0$ on the right piece of the timelike component of $\partial\mathscr{B}_{\bar{x}}$. Therefore, only the integral on the left timelike piece contributes to the right hand side of the previous equation. From this and equation (A30), together with the assumption that the region $\mathscr{B}_{\bar{x}}$ is infinitesimally thin around the wall, we get

$$-4\pi\int d\tau dy dz\bar{x}(\sigma_a Y^a) = \int d\tau dy dz\bar{x}(F_{ab}X^aY^b), \tag{A41}$$

where we used that the volume density is $\bar{x}d\tau dydz$ and that the normal (oriented for the Gauss theorem is) $n^a = X^a$; however, the conventional inner pointing normal of the box N^a. Assuming that the region is infinitesimal in all directions the previous identity implies $F_{ab}X^aY^b = -4\pi\sigma_a Y^a$ which in terms of N^a reads

$$F_{ab}N^aY^b = 4\pi\sigma_a Y^a. \tag{A42}$$

The same logic applied to the current p_a^Z implies

$$F_{ab}N^aZ^b = 4\pi\sigma_a Z^a. \tag{A43}$$

A moment of reflection shows that the argument is also true for the current p_a^T. Now the top and bottom contributions vanish because the normal there $n^a \propto T^a$ and thus $F_{ab}n^aT^a = 0$ because of the skew symmetry of F_{ab}. Therefore, we also have

$$F_{ab}N^aT^b = 4\pi\sigma_a T^a. \tag{A44}$$

The most general F_{ab} would be of the form

$$F_{ab} = f_{TN}T_{[a}N_{b]} + f_{TY}T_{[a}Y_{b]} + f_{TZ}T_{[a}Z_{b]} + f_{YN}Y_{[a}N_{b]} + f_{ZN}Z_{[a}N_{b]} + f_{YZ}Y_{[a}Z_{b]}. \tag{A45}$$

The fact that the electric field $E_a^{\rm box}$ is proportional to N^a due to the presence of the wall implies that $f_{TY} = f_{TZ} = 0$. As the normal component of the magnetic field must vanish due to the present of the conducting wall we also have $f_{YZ} = 0$. Thus

$$F_{ab} = f_{TN}T_{[a}N_{b]} + f_{YN}Y_{[a}N_{b]} + f_{ZN}Z_{[a}N_{b]}. \tag{A46}$$

Equations (A42), (A43), and (A44) fix the last three components. The solution is

$$\boxed{F_{ab} = 8\pi N_{[a}\sigma_{b]}}, \tag{A47}$$

which is the same as (A33). One can easily check that the same solution follows from the same argument applied to regions $\mathscr{B}_y$ and $\mathscr{B}_z$ adapted to the world sheets of the other walls of the box.

Appendix B: Work done by the walls (Maxwell case)

We have seen that the energy content of the box, as measured in the lab frame, depends on τ. This is due to the action of an external agent that is accelerating the box of radiation. The change in the energy $E(\tau)$ is thus related to the work done by the external agent on the box. This is associated with the failure for the current j_a to be conserved: in our idealization of the accelerating box, the external agent acts upon the electromagnetic field via the boundary charges that impose the box boundary conditions and source the divergence of j_a (see (A10)). The dynamical contribution of these charges is feeding energy into the system. In fact, from the Gauss law, now applied to j_a in the region of interest (Figure 2) we get

$$\begin{aligned} E(\tau) - E(0) = \Delta W &\equiv \int_R \nabla^a j_a^{\rm lab} - \int_{\partial R - \Sigma_1 - \Sigma_2} j_a^{\rm lab} N^a \\ &= 2\pi \int_{\partial R - \Sigma_1 - \Sigma_2} (\sigma\cdot\sigma) N_b \xi^b_{\rm lab} \end{aligned} \tag{B1}$$

where we have used (A35) and (A36) and the Gauss theorem where the bulk integral involves the integration of the $\delta_{\rm box}$ distribution whose support is at the boundary of R. Now from (10) we observe that $N_b\xi^b_{\rm lab} = 0$ on any parts of the boundary where $\partial_{\bar{x}}$ is tangent to the boundary. At the bottom and at the top we have $N_b\xi^b_{\rm lab} = \mp\gamma(v)v$ (where for simplicity we are assuming that the box is a cube with walls defined by $\bar{x}$, y, and z equal constant). Therefore, using this in the last line of (B1) we get

$$E(\tau) - E(0) = \left(\int_{\rm top} - \int_{\rm bottom}\right) 2\pi\gamma v(\sigma\cdot\sigma). \tag{B2}$$

On the other hand, the pressure on the top/bottom is given by

$$P_{\bar{x}} \equiv T_{ab}\partial^a_{\bar{x}}\partial^b_{\bar{x}} = T_{\bar{x}\bar{x}} = 2\pi(\sigma\cdot\sigma). \tag{B3}$$

So we find the following expression for the work

$$E(\tau)-E(0) = \left(\int_{\rm top}-\int_{\rm bottom}\right) P_{\bar{x}}\bar{x}\gamma v d\tau dydz \tag{B4}$$

$$= \int \sinh\tau d\tau \left(\bar{x}\int P_{\bar{x}}dydz\right)\Big|_{\rm bottom}^{\rm top}. \tag{B5}$$

This expression tells us that the origin of the inertia of the box (its resistance to acceleration encoded in the mass (32)) is the difference of the radiation pressure of the electromagnetic field on the walls between the top and the bottom of the box. In order to accelerate the box, an external agent must impose an external force to compensate for the radiation pressure of the confined radiation. Its infinitesimal version is

$$\frac{dE}{d\tau} = \gamma v \left(\bar{x}\int P_{\bar{x}}dydz\right)\Big|_{\rm bottom}^{\rm top}. \tag{B6}$$

Let us define

$$F^{\bar{x}} = \int P_{\bar{x}}dydz \tag{B7}$$

In order to better interpret the previous result let us introduce the proper time measured at the center of the box $\mathrm{T_c}=\bar{x}_{\rm c}\tau$, and recall that we denote by L the legth of the box. With this the previous equation becomes

$$\frac{dE}{d\mathrm{T}_c} = \gamma v\left(\left(1+\frac{L}{2\bar{x}_{\rm c}}\right)F^{\bar{x}}_{\rm top}-\left(1-\frac{L}{2\bar{x}_{\rm c}}\right)F^{\bar{x}}_{\rm bottom}\right), \tag{B8}$$

and using that the magnitude of the acceleration of the center of the box is (according to (17)) $|a|=c^2/\bar{x}_c$ we arrive at

$$\frac{dE}{d\mathrm{T}_c} = \gamma v\left(F^{\bar{x}}_{\rm top}-F^{\bar{x}}_{\rm bottom}\right)+\frac{|a|L}{2c^2}\gamma v\left(F^{\bar{x}}_{\rm top}+F^{\bar{x}}_{\rm bottom}\right). \tag{B9}$$

Let us define $F_{\rm net}\equiv\left(F^{\bar{x}}_{\rm top}-F^{\bar{x}}_{\rm bottom}\right)$ for what follows. If both the length and acceleration of the box are small in the sense that by the time light travels the distance L the velocity increase of the box due to the acceleration is much smaller than c then

$$\frac{|a|L}{2c^2}\ll 1. \tag{B10}$$

For such small-size/small-acceleration boxes (those that model well a composite particle) we recover the usual relativistic law

$$\frac{\mathrm{d}E}{\mathrm{dT_c}} = \gamma v F_{\rm net}. \tag{B11}$$

or in covariant notation $u^a_{\rm c}\nabla_a E = F^a_{\rm net}\xi^{\rm lab}_a$. The previous equation implies the familiar second Newton equation in the instantaneous rest frame

$$ma = F_{\rm net} \tag{B12}$$

where m is given by $E(0)/c^2$ as given in (29). This shows that the physical origin of mass can be traced to the inertia produced by the difference of radiation pressure between the top and the bottom of the box (in analogy with the heuristic simplistic picture given in terms of the bouncing photon in Section II).

Appendix C: Work done by the walls (scalar field case)

In this section we repeat the derivation of the previous one but in the case of the massless scalar field. In the lab frame $\xi^a_{\rm lab}=\partial^a_t$ the associated energy-momentum current is

$$j^{\rm lab}_a \equiv -T_{ab}\xi^b_{\rm lab} \tag{C1}$$

which, as in the electromagnetic case, is not conserved due to the contributions of the boundary degrees of freedom. One has

$$\nabla_a j^a_{\rm lab} = -\sigma\, u^b \nabla_b \phi \delta_{\rm box} \tag{C2}$$

where we used (5). The energy momentum tensor (34) at the walls is therefore

$$T_{ab}|_{\rm box} = \sigma^2 \left(N_a N_b - \frac{1}{2} g_{ab} \right) \tag{C3}$$

Now from (35) we have

$$\nabla^a T_{ab} = \sigma^2 N_b \delta_{\rm box}, \tag{C4}$$

from (C1)

$$j_a N^a|_{\rm box} = -\frac{1}{2}\sigma^2 N_b u^b, \tag{C5}$$

and from (C2)

$$\nabla^a j_a = -\sigma^2 N_b u^b \delta_{\rm box}. \tag{C6}$$

Finally, notice that the pressure at the walls is given by

$$P_N \equiv T_{ab} N^a N^b = \frac{1}{2}\sigma^2. \tag{C7}$$

As in the electromagnetic case, equations (C4), (C5), (C6), and (C7) and the same line of argument of Section B lead to

$$\frac{dE}{d\tau} = \gamma v \left(\bar{x} \int P_{\bar{x}} dydz \right) \Big|_{\rm bottom}^{\rm top}. \tag{C8}$$

The same conclusions as for the Maxwell case follow from here.

---

[1] A. Einstein, *Relativity: the Special and the General theory.* 100th anniversary ed., 2015. With commentaries and background material by Hanoch Gutfreund and Jürgen Renn.

[2] A. Einstein, "Does the Inertia of a Body Depend Upon its Energy-Content," Annalen der Physik **323** (March, 1905) 639–641.

[3] J. D. Jackson, *Classical electrodynamics; 2nd ed.* Wiley, New York, NY, 1975.

[4] S. Durr *et al.*, "Ab-Initio Determination of Light Hadron Masses," Science **322** (2008) 1224–1227, arXiv:0906.3599.

[5] E. Schrödinger, *Über die kräftefreie Bewegung in der relativistischen Quantenmechanik.* Akademie der wissenschaften in kommission bei W. de Gruyter u. Company, 1930.

[6] W. Rindler, *Essential Relativity: Special, General, and Cosmological.* Springer New York, 2013.

[7] R. Wald, *General Relativity.* University of Chicago Press, Chicago, 1984.

[8] W. Rindler, "Hyperbolic Motion in Curved Space Time," Phys. Rev. **119** (Sep, 1960) 2082–2089.

[9] K. S. Thorne and D. MacDonald, "Electrodynamics in curved spacetime: 3 + 1 formulation," Monthly Notices of the Royal Astronomical Society **198** (02, 1982) 339–343, arXiv:https://academic.oup.com/mnras/article-pdf/198/2/339/9402846/mnras198-0339.pdf.