Open Science Framework DOI:10.31219/osf.io/***

Notes on energy and momentum

P.G.L. Porta Mana
Western Norway University of Applied Sciences <pgl@portamana.org>

22 December 2024; updated 25 December 2024 [draft]

***

1 Questions about momentum and energy

The notions of momentum and energy in Newtonian physics present many peculiarities, at least to the eye of a student who's learning about them. Consider for instance the following questions, some of which have indeed been asked by students, including me:

(a) How should *total energy* be defined? Is it just the sum of internal and kinetic energies? or should it also include gravitational potential energy?

(b) Why is *internal energy* the same in two reference frames, whereas *kinetic energy* differs?

(c) Why does force, which is related to change of momentum, appear in the laws for the change of energy (in the formula for "work")? Why doesn't change of energy appear, vice versa, in the law for change of momentum?

(d) What is the *law of transformation for momentum* between reference frames? For instance, if we know that the components of momentum in an inertial frame are $[0, 1, 2]\,\mathrm{N\,s}$, then what are the components in another inertial frame with constant velocity $\boldsymbol{v}$ with respect to the first?

Tentative answers to some of these questions may lead to embarrassing further questions.

Take question (a) for example. A possible answer is that the definition of "total energy" is arbitrary. For an object near Earth's surface, we may avoid speaking of gravitational potential energy if we include the work done by gravitational forces in the balance of energy; or vice versa we may include a gravitational potential energy in the total energy, avoiding the inclusion of work by gravitational forces in the energy balance. To this explanation the student may ask why we have such definition freedom for energy, but not an analogous one for momentum.

Or take question (d). A possible answer is that in order to know the momentum in another frame we need to know both the mass and the velocity of an object. At this answer the student may have the following questions: What in the case of an electromagnetic field, where there's no mass or velocity? And why do we need such extra information for the transformation of momentum, when we don't need extra information for the transformation of velocity or of mass?

There are different perspectives from which one can try to answer questions like these. One can take a historical, rather than physical, point of view. Or one may say that it's just a matter of definitions. The literature also offers more physical answers for some of these questions, based for instance on symmetry, or on variational principles, or on the Newton-Cartan theory. As an example, some work of Šilhavý[1], derives the notions of mass, momentum, and kinetic energy from Galileian invariance.

Here we show how these questions receive answers from the point of view of general relativity.

2 Energy-momentum tensor

2.1 Representation

Energy, momentum, flux of energy, and force, or more generally flux of momentum, are all components of a single object, the *energy-momentum tensor*, also called 'energy-momentum-stress tensor', with permutations, or 'stress-energy tensor', or 'four-stress', or also 'mass tensor'. It can be represented by a 4-by-4 matrix. Roughly speaking, the entries in the

[1] Šilhavý 1989; 1992; see also Serrin 1998.

matrix can be related as follows with energy and momentum:

$$\begin{bmatrix} \text{energy} & \text{momentum} \\ \text{energy flux} & \text{force (momentum flux)} \end{bmatrix} . \tag{1}$$

The "roughly speaking" will be made precise soon.

The fact that energy, momentum, and their fluxes are components of a simple object, already gives a partial answer to question (c). We know in the simpler examples of vectors that their components get mixed in a change of frame. This makes it plausible that forces, for instance, should appear in the energy flux upon a change of frame. Also question (d) starts to receive an answer. Generally all components of an object are needed to find the new ones in another frame. The three components of momentum are not really a vector by themselves, but are partial components of something larger; this is why they are not enough, by themselves, for finding the momentum components in a new frame. This also hints at the fact that *mass* or something related to it should be one of the additional components.

The energy-momentum tensor can be represented in several equivalent ways, differing in their contravariant or covariant character. Here we discuss two: one that is common in textbook, and one that has a more immediate physical meaning and is more often used for numerical computations. The first is by a contra-contra-variant tensor

$$\boldsymbol{T} = T^{\mu\nu}\, \boldsymbol{e}_\mu \otimes \boldsymbol{e}_\nu \tag{2}$$

and the second by a tensor density

$$\boldsymbol{\mathcal{T}} = \mathcal{T}^{\mu}{}_{\nu}\, \mathrm{d}^3X_\mu \otimes \mathrm{d}x^\nu \tag{3}$$

The components of the two representations are related by

$$\mathcal{T}^{\mu}{}_{\nu} = \frac{\sqrt{|g|}}{c}\, T^{\mu\alpha}\, g_{\alpha\nu} \tag{4}$$

where g is the determinant of the metric $\boldsymbol{g}$.

The component $\mathcal{T}^{\mu}{}_{\nu}$ expresses, roughly speaking, the flow of x^{ν}-momentum *per unit coordinates*, across a volume (hypersurface in spacetime) of constant x^{μ}. Again the "roughly speaking" will be made precise soon.

In the expressions above, $\mathrm{d}x^{\mu}$ are the differentials of the coordinates x^{mu} and form a basis for covector fields, and $\boldsymbol{e}_{\mu}$ are the dual vectors and form a basis for vector fields. The d^3X_{μ} are volume forms defined as follows:[2]

$$\mathrm{d}^3X_0 \coloneqq \mathrm{d}\tilde{x}^1\mathrm{d}x^2\mathrm{d}x^3 \coloneqq \mathrm{d}\tilde{x}^1 \wedge \mathrm{d}x^2 \wedge \mathrm{d}x^3$$

$$\mathrm{d}^3X_1 \coloneqq -\mathrm{d}\tilde{x}^0\mathrm{d}x^2\mathrm{d}x^3 \qquad \mathrm{d}^3X_2 \coloneqq -\mathrm{d}\tilde{x}^0\mathrm{d}x^3\mathrm{d}x^1 \qquad \mathrm{d}^3X_3 \coloneqq -\mathrm{d}\tilde{x}^0\mathrm{d}x^1\mathrm{d}x^2 \tag{5}$$

that is, they are a basis for 3-covector fields. They can be integrated over three-dimensional regions of spacetime and therefore are the main building block for the definition of the total amount of any quantity in a spatial region, and of the flow of any quantity through a surface during a time lapse. The tilde $\tilde{}$, called an outer-oriented scalar, indicates that an outer orientation is chosen for these 3-covectors, instead of an inner orientation. For instance, the 3-covector $\mathrm{d}\tilde{x}^1\mathrm{d}x^2\mathrm{d}x^3$ does not have a screw-sense $x^1x^2x^3$, but an orientation along the complementary direction x^0, towards positive x^0. This complementary orientation is indicated by the subscript on "d^3X". Compare with the similar notation in Misner et al. 2017 ch. 2 Box 5.4, who however define the volume forms slightly differently; or in Weinberg 1972 § 4.11.

2.2 Equations for the energy-momentum tensor

The energy-momentum tensor satisfies two fundamental equations or constraints, as a consequence of the Einstein equations. The first is

$$\mathrm{D}\boldsymbol{\mathcal{T}} = 0 \qquad \text{or in components} \quad \partial_{\mu}\mathcal{T}^{\mu}{}_{\alpha} - \mathcal{T}^{\mu}{}_{\nu}\,\Gamma^{\nu}_{\mu\alpha} = 0 \tag{6}$$

where D is the covariant exterior derivative. The second is, directly in components,

$$\mathcal{T}^{\mu}{}_{\alpha}\,g^{\alpha\nu} = \mathcal{T}^{\nu}{}_{\alpha}\,g^{\alpha\mu}\,. \tag{7}$$

The first equation is often presented as saying that "the divergence of the energy-momentum tensor is zero". But we must keep in mind that

[2] notation similar to Gotay & Marsden 1992 § 2 p. 371.

this "divergence" does *not* satisfy a Stokes theorem. The fundamental problem underlying this is that *the energy-momentum tensor cannot be integrated over a spacetime region*, be it four- or three-dimensional. So it doesn't make sense to speak of the "total" energy-momentum in a region. The same is true for the divergence of the energy-momentum tensor. This is a basic point in differential geometry: tensors at different points cannot be compared, added, and so on, in any standard way; there are different inequivalent ways to realize such comparisons and sums.

Then how come that we usually speak of the total energy or momentum in a region, and of their fluxes through surfaces?

We shall now see how the energy-momentum tensor can be used to define energy and momenta that satisfy particular balance laws, thanks to equations (6)–(7). The definition, however, depends on extra fields; so it's important to keep in mind is that *in general there is no unique, canonical definition* of such energy and momentum. The energy and momentum used in Newtonian mechanics are just particular choices, related to the approximations involved in Newtonian mechanics.

2.3 Definitions of energy and momentum

We can choose a non-zero vector field $\boldsymbol{V}$ over a four-dimensional spacetime region, which we shall call *reference field*, and construct an associated 4-current[3]

$$\boldsymbol{\mathcal{J}} \coloneqq \boldsymbol{\mathcal{T}} \cdot \boldsymbol{V} \qquad \text{in components} \quad \mathcal{J}^{\mu}\, \mathrm{d}^3 X_{\mu} = (\mathcal{T}^{\mu}{}_{\alpha} V^{\alpha})\, \mathrm{d}^3 X_{\mu} \,. \tag{8}$$

This current is a 3-covector, so it can be integrated over any three-dimensional region of spacetime. What's more, thanks to eqs (6) and (7) this current satisfies the following balance law:

$$\mathrm{d}\boldsymbol{\mathcal{J}} = \tfrac{1}{2}\operatorname{tr}\bigl(\boldsymbol{T} \cdot \mathrm{L}_{\boldsymbol{V}}\boldsymbol{g}\bigr)\, \frac{\sqrt{|g|}}{c} \mathrm{d}^4 X \,. \tag{9}$$

Stokes's theorem holds for this current: for any four-dimensional region R where $\boldsymbol{V}$ is defined,

$$\begin{aligned} \int_{\partial R} \boldsymbol{\mathcal{J}} &= \int_R \mathrm{d}\boldsymbol{\mathcal{J}} \\ &= \int_R \tfrac{1}{2}\operatorname{tr}\bigl(\boldsymbol{T} \cdot \mathrm{L}_{\boldsymbol{V}}\boldsymbol{g}\bigr)\, \frac{\sqrt{|g|}}{c} \mathrm{d}^4 X \,. \end{aligned} \tag{10}$$

[3] Gotay & Marsden 1992; Hawking & Ellis 1994 § 3.2 p. 62; Choquet-Bruhat & DeWitt-Morette 2000 § II.7.III p. 87; see also the discussion in van Dantzig 1934 part 4 § 1.

If the reference field $\boldsymbol{V}$ is a Killing vector in the spacetime region, so that $\mathrm{L}_{\boldsymbol{V}}\boldsymbol{g} = 0$ there, then the current $\boldsymbol{\mathcal{J}}$ is conserved there.

The ten equations (6)–(7) satisfied by the energy-momentum tensor can be alternatively formulated as balance laws satisfied by ten currents associated to ten independent reference fields. This is how the balance of energy, of the three components of momentum, of three components of angular momentum, and of three components of boost-momentum come about.

We speak of energy for the 4-current associated with a timelike reference field. We speak of three components of momentum for the 4-currents associated with three spacelike reference fields. This view of energy and momentum as defined with respect to "reference vector fields" encompasses other common points of view, especially about energy. An example is the virtual-work point of view: the vector $\boldsymbol{V}$ in this case represents a field of virtual displacements. Another clear example is the point of view connected with symmetries and Noether's theorem.

Clearly we have infinite choices of vector fields to associate energy and momentum with. We could also define currents that are somehow hybrids of energy and momentum, or "momentum" whose components don't really form a vector. How to choose such reference fields?

Suppose we have a coordinate system (t, x, y, z), with timelike t and spacelike x, y, z. Then there are several natural choices of reference fields with which we can associate 4-currents in the manner above.

One choice are the basis vector fields

$$\boldsymbol{e}_t \qquad \boldsymbol{e}_x \qquad \boldsymbol{e}_y \qquad \boldsymbol{e}_z\ , \tag{11}$$

energy being associated with the first, and momentum with the last three. For spacelike coordinates that have angular dimension, the associated momentum has the character of an angular momentum. Note that these vector fields may not be orthonormal.

Another natural choice are the vectors above, but appropriately normalized:

$$\frac{c}{\sqrt{|g_{tt}|}}\boldsymbol{e}_t \qquad \frac{1}{\sqrt{|g_{xx}|}}\boldsymbol{e}_x \qquad \frac{1}{\sqrt{|g_{yy}|}}\boldsymbol{e}_y \qquad \frac{1}{\sqrt{|g_{zz}|}}\boldsymbol{e}_z\ , \tag{12}$$

because $\boldsymbol{e}_t \cdot \boldsymbol{g} \cdot \boldsymbol{e}_t = g_{tt}$ and so on. Normalizing them makes sense in that the displacements represented by these vectors are in terms of unit physical time and unit physical length.

Two other natural choices are the vectors obtained from the basis covectors:

$$c\,\boldsymbol{g}^{-1}\cdot \mathrm{d}t \qquad \boldsymbol{g}^{-1}\cdot \mathrm{d}x \qquad \boldsymbol{g}^{-1}\cdot \mathrm{d}y \qquad \boldsymbol{g}^{-1}\cdot \mathrm{d}z\;, \tag{13}$$

as well as their normalized counterparts analogous to (12).

Further natural choices of reference fields are possible, unrelated to the coordinate system. One may choose, for instance, Killing vector fields. Or one may choose vectors fields which are orthonormal to one another, a so-called tetrad field. In the presence of matter there is a four-velocity field $\boldsymbol{U}_\mathrm{m}$, and we can choose this timelike field as the reference one to associate an energy with.

In low-curvature regions typical of the Newtonian approximation, and with approximately inertial coordinates, it turns out that all choices above for the *spacelike* reference fields lead to approximately the same three currents. This is why there seems to be a unique definition of momentum, and why we can consider the x-, y-, z-momenta as components of a vector:

$$\boldsymbol{\mathcal{T}}\cdot\boldsymbol{e}_x \approx \frac{1}{\sqrt{|g_{xx}|}}\boldsymbol{\mathcal{T}}\cdot\boldsymbol{e}_x \approx \boldsymbol{\mathcal{T}}\cdot\boldsymbol{g}^{-1}\cdot\mathrm{d}x \approx \mathcal{T}^{\mu}{}_{x}\,\mathrm{d}^3X_\mu =: \boldsymbol{\mathcal{P}}_x \tag{14}$$

and similarly for y and z.

The different choices of *timelike* reference field, however, lead to slightly different associated currents. This is why there seems to be some freedom in the definition of energy. In particular:

- The energy four-current

$$\boldsymbol{\mathcal{E}}_\mathrm{i} := \boldsymbol{\mathcal{T}}\cdot\boldsymbol{U}_\mathrm{m} \tag{15}$$

  associated with the matter four-velocity $\boldsymbol{U}_\mathrm{m}$ is what we call *internal energy-mass*. Note that this reference field is unrelated to the coordinates. This is why internal energy is an invariant in the Newtonian approximation and even in the exact case.

- The energy four-current

$$\boldsymbol{\mathcal{E}}_\mathrm{i} + \boldsymbol{\mathcal{E}}_\mathrm{k} := \frac{c}{\sqrt{|g_{tt}|}}\boldsymbol{\mathcal{T}}\cdot\boldsymbol{e}_t \tag{16}$$

  associated with the normalized basis field $\frac{c}{\sqrt{|g_{tt}|}}\boldsymbol{e}_t$ is represented as the sum of internal energy above and of *kinetic energy*. Kinetic energy is therefore associated with the difference between the two reference fields above:

$$\boldsymbol{\mathcal{E}}_{\mathrm{k}} \coloneqq \boldsymbol{\mathcal{T}} \cdot \left(\frac{c}{\sqrt{|g_{tt}|}}\boldsymbol{e}_t - \boldsymbol{U}_{\mathrm{m}}\right) . \tag{17}$$

Note that this difference is a spacelike vector field, therefore kinetic energy has something akin to momentum, so to speak.

- The energy four-current

$$\boldsymbol{\mathcal{E}}_{\mathrm{i}} + \boldsymbol{\mathcal{E}}_{\mathrm{k}} + \boldsymbol{\mathcal{E}}_{\mathrm{p}} \coloneqq \boldsymbol{\mathcal{T}} \cdot \boldsymbol{e}_t \equiv \mathcal{T}^{\mu}{}_{t}\, \mathrm{d}^3 X_{\mu} \ . \tag{18}$$

associated with the basis field $\boldsymbol{e}_t$ is represented as the sum of internal, kinetic, and *gravitational potential* energy. Potential gravitational energy is therefore associated with the difference between two reference fields:

$$\boldsymbol{\mathcal{E}}_{\mathrm{p}} \coloneqq \boldsymbol{\mathcal{T}} \cdot \left(1 - \frac{c}{\sqrt{|g_{tt}|}}\right) \boldsymbol{e}_t \ . \tag{19}$$

In Schwarzschild isotropic coordinates typical of the barycentric or geocentric celestial reference systems (BCRS, GCRS) used in astronomy[4], the vector field $\boldsymbol{e}_t$ is a Killing vector field; therefore the associated energy $\boldsymbol{\mathcal{E}}_{\mathrm{i}} + \boldsymbol{\mathcal{E}}_{\mathrm{k}} + \boldsymbol{\mathcal{E}}_{\mathrm{p}}$ is conserved, not just balanced; that is, the right side of eq. (9) is zero.

Thus we see where the arbitrariness in the definition of total energy in Newtonian mechanics comes from: it corresponds to the arbitrariness in choosing a timelike reference field for defining energy.

Bibliography

("de *X*" is listed under D, "van *X*" under V, and so on, regardless of national conventions.)

Buttazzo, G., Galdi, G. P., Lanconelli, E., Pucci, P., eds. (1998): *Nonlinear Analysis and Continuum Mechanics: Papers for the 65th Birthday of James Serrin*. (Springer, New York).

Choquet-Bruhat, Y., DeWitt-Morette, C. (2000): *Analysis, Manifolds and Physics. Part II: 92 Applications*, rev. enl. ed. (Elsevier, Amsterdam). doi:10.1016/B978-0-444-50473-9.X5000-3. First publ. 1989.

Gotay, M. J., Marsden, J. E. (1992): *Stress-energy-momentum tensors and the Belinfante-Rosenfeld formula*. Contemp. Math. **132**, 367–392. https://www.cds.caltech.edu/~marsden/bib/1992/05-GoMa1992/, doi:10.1090/conm/132.

[4] Kaplan 2005; Soffel et al. 2003; Petit & Wolf 2005; Soffel & Langhans 2013.

Hawking, S. W., Ellis, G. F. R. (1994): *The Large Scale Structure of Space-Time*, repr. (Cambridge University Press, Cambridge). doi:10.1017/CBO9780511524646, doi:10.1017/9781009253161. First publ. 1973.

Kaplan, G. H. (2005): *The IAU resolutions on astronomical reference systems, time scales, and Earth rotation models*. Tech. rep. 179. (US Naval Observatory, Washington). https://aa.usno.navy.mil/publications/Circular_179, arXiv doi:10.48550/arXiv.astro-ph/0602086.

Misner, C. W., Thorne, K. S., Wheeler, J. A. (2017): *Gravitation*, repr. (Princeton University Press, Princeton and Oxford). With a new foreword by David I. Kaiser and a new preface by Charles W. Misner and Kip S. Thorne. First publ. 1970. https://archive.org/details/GravitationMisnerThorneWheeler.

Petit, G., Wolf, P. (2005): *Relativistic theory for time comparisons: a review*. Metrologia **42**³, S138–S144. doi:10.1088/0026-1394/42/3/S14, http://geodesy.unr.edu/hanspeterplag/library/geodesy/time/met5_3_S14.pdf.

Serrin, J. (1998): *Space, time and energy*. In: Buttazzo, Galdi, Lanconelli, Pucci (1998): ch. 11:139–144. First publ. in Italian 1995.

Šilhavý, M. (1989): *Mass, internal energy, and Cauchy's equations in frame-indifferent thermodynamics*. Arch. Ration. Mech. Anal. **107**¹, 1–22.

Šilhavý, M. (1992): *Energy principles and the equations of motion in Galilean thermomechanics*. Czechoslov. J. Phys. **42**⁴, 363–374.

Soffel, M., Klioner, S. A., Petit, G., Wolf, P., Kopeikin, S. M., Bretagnon, P., Brumberg, V. A., Capitaine, N., et al. (2003): *The IAU 2000 resolutions for astrometry, celestial mechanics, and metrology in the relativistic framework: explanatory supplement*. Astron. J. **126**⁶, 2687–2706. doi:10.1086/378162.

Soffel, M., Langhans, R. (2013): *Space-Time Reference Systems*. (Springer, Berlin). doi:10.1007/978-3-642-30226-8.

van Dantzig, D. (1934): *Electromagnetism, independent of metrical geometry. 1. The foundations. 2. Variational principles and further generalisation of the theory. 3. Mass and motion. 4. Momentum and energy; waves*. Proc. Acad. Sci. Amsterdam **37**⁸⁻¹⁰, 521–525, 526–531, 643–652, 825–836. See also van Dantzig (1936). https://dwc.knaw.nl/DL/publications/PU00016602.pdf, https://dwc.knaw.nl/DL/publications/PU00016603.pdf, https://dwc.knaw.nl/DL/publications/PU00016620.pdf, https://dwc.knaw.nl/DL/publications/PU00016648.pdf.

van Dantzig, D. (1936): *Electromagnetism, independent of metrical geometry. 5. Quantum-theoretical commutability-relations for light-waves*. Proc. Acad. Sci. Amsterdam **39**¹, 126–131. See also van Dantzig (1934). https://dwc.knaw.nl/DL/publications/PU00016835.pdf.

Weinberg, S. (1972): *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. (Wiley, New York).