Energy and Gravitational Waves

In flat Minkowski spacetime, and using standard inertial coordinates x^0, x^1, x^2, x^3 such that the metric takes the form

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2\,,$$

we can define the total 4-momentum of a system with stess-energy tensor $T^{\mu\nu}$ to be given by

$$P^\mu(t) = \frac{1}{c}\int T^{\mu 0}(ct, \boldsymbol{x})\, d^3\boldsymbol{x}\,.$$

Since the stress energy tensor obeys the continuity equation

$$\partial_\nu T^{\mu\nu} = \frac{1}{c}\frac{\partial T^{\mu 0}}{\partial t} + \frac{\partial T^{\mu i}}{\partial x^i} = 0\,,$$

the total 4-momentum must be conserved. This is because

$$\frac{dP^\mu}{dt} = \frac{1}{c}\int \frac{\partial T^{\mu 0}}{\partial t}\, d^3\boldsymbol{x} = -\int \frac{\partial T^{\mu i}}{\partial x^i}\, d^3\boldsymbol{x}\,,$$

and this final integral can be transformed into one over an infinitely distant surface via Gauss' theorem, which implies it must be equal to zero for any localised stress-energy distribution.

We seek to adapt this definition to a spacetime with a weak gravitational field in order to develop the notion of energy carried by gravitational waves. By definition, a gravitational field is weak if there exists a set of coordinates x^0, x^1, x^2, x^3 in which the metric takes the form

$$ds^2 = (dx^0)^2 - (dx^1)^2 - (dx^2)^2 - (dx^3)^2 + h_{\mu\nu}dx^\mu dx^\nu$$

with $h_{\mu\nu}$ being small. These coordinates are certainly not unique, and in general the $h_{\mu\nu}$ are not well defined due to the presence of gauge transformations. As a consequence of this fact it will essentially be impossible for us to formulate the discussion in a coordinate independent manner, and so the notion of energy we construct will be inherently dependent on the particular choice of coordinates we make. However arguably this shouldn't be surprising as this is also true in classical Newtonian mechanics. The gauge transformations of general relativity correspond (roughly speaking) to the ability of Newtonian mechanics to transfer into a non-inertial frame via the introduction of fictitious forces. These fictitious forces can come with associated fictitious potentials, and so lead to different expressions for energy than would be obtained in an inertial frame.

In a "nearly inertial frame" described by the weak field metric above, we can define the 4-momentum of matter to be given by

$$P^\mu_{\text{mat}}(t) = \frac{1}{c}\int T^{\mu 0}(ct, \boldsymbol{x})\, d^3\boldsymbol{x}\,.$$

Note that this is simply a definition, we are not claiming anything about the physics of this quantity, and the fact that we are calling it a 4-momentum at all is simply due to its visual similarity with the special relativistic case. The motivating idea behind this construction is that, if we equipped an observer with our chosen coordinate system, and did not inform them that they were in a curved spacetime, this is what they would believe the total 4-momentum of the matter present to be. However it must be stressed that P^μ_{mat} is not actually a 4-vector, it is simply an object with four components which we can pretend is one when the mathematicians aren't looking.

Importantly, since in curved spacetime it is $\nabla_\nu T^{\mu\nu}$ and not $\partial_\nu T^{\mu\nu}$ which is zero, P^μ_{mat} will not be conserved in the same way as it is in special relativity. From the perspective of our naive observer blissfully unaware of the non-Minkowski nature of the spacetime they inhabit, this non conservation of P^μ_{mat}, which they believe really is a 4-momentum, would suggest that there must be some other source of 4-momentum that they are not accounting for. That is to say that in order to balance their books, this observer would be forced to conclude that the total 4-momentum in spacetime is given by

$$P^\mu_{\text{tot}} = P^\mu_{\text{mat}}(t) + P^\mu_{\text{grav}}(t) = \text{constant}\,,$$

where the suggestively named P^μ_{grav} is a source of 4-momentum which is somehow not encapsulated by the stress-energy tensor. It is this quantity that we decide to call the energy and momentum carried by the gravitational field. Since the four numbers that make up P^μ_{grav} have significantly fewer degrees of freedom than a set of continuous functions, we will always be able to find (an infinite number of) fields $t^{\mu\nu}$ such that

$$P^\mu_{\text{grav}}(t) = \frac{1}{c}\int t^{\mu 0}(ct,\boldsymbol{x})\,d^3\boldsymbol{x}\,,$$

meaning that

$$P^\mu_{\text{tot}} = \frac{1}{c}\int [T^{\mu 0}(ct,\boldsymbol{x}) + t^{\mu 0}(ct,\boldsymbol{x})]d^3\boldsymbol{x} = \text{constant}\,.$$

From our discussion of special relativity, it follows that one easy way of achieving this would be to choose $t^{\mu\nu}$ such that

$$\partial_\nu(T^{\mu\nu} + t^{\mu\nu}) = 0\,.$$

This would have the added benefit that our naive observer would be able to interpret $t^{\mu\nu}$ as the stress-energy "tensor" of the gravitational field and have 4-momentum conservation hold in a local sense as well as globally. Now if $t^{\mu\nu}$ represents the stress-energy of the gravitational field we might hope that it is expressible in terms of only the metric. One way we could go looking for such a $t^{\mu\nu}$ is to first construct an auxiliary field $H^{\mu\nu}$ from the metric, such that it obeys

$$\partial_\nu H^{\mu\nu} = 0\,.$$

We could then define $t^{\mu\nu}$ by the equation

$$t^{\mu\nu} = \frac{c^4}{8\pi G}\left(G^{\mu\nu} - H^{\mu\nu}\right).$$

It would then be guaranteed by Einstein's field equations

$$G^{\mu\nu} = -\frac{8\pi G}{c^4}T^{\mu\nu}\,,$$

that $\partial_\nu(T^{\mu\nu} + t^{\mu\nu}) = 0$. By a happy coincidence we already know a suitable candidate $H^{\mu\nu}$ which possesses this property. The first order approximation to the Einstein tensor

$$H^{\mu\nu} = \frac{1}{2}\left(\partial_\sigma\partial^\sigma\bar{h}^{\mu\nu} + \eta^{\mu\nu}\partial_\sigma\partial_\tau\bar{h}^{\sigma\tau} - \partial_\sigma\partial^\mu\bar{h}^{\sigma\nu} - \partial_\sigma\partial^\nu\bar{h}^{\mu\sigma}\right) \quad \text{where} \quad \bar{h}^{\mu\nu} = h^{\mu\nu} - \eta^{\mu\nu}\frac{h^{\sigma\tau}\eta_{\sigma\tau}}{2}\,,$$

can be seen to satisfy $\partial_\nu H^{\mu\nu} = 0$ as a result of the symmetry of mixed partial derivatives. Of course, any multiple of this would also work, but this particular choice of normalisation has the benefit that $t^{\mu\nu}$ ends up being second order in h, which aligns with our usual notion of energy as depending quadratically on the amplitude of a wave.

The principal things to note about $t^{\mu\nu}$ is that it is neither a tensor, nor gauge invariant, and is thus fundamentally and inextricably tied to the coordinate system we used to calculate it. Nonetheless, so long as we calculate all energies in this coordinate system the fact that $\partial_\nu(T^{\mu\nu} + t^{\mu\nu}) = 0$ does allow us to interpret it as a stress energy of the gravitational field. For example, we can treat t^{0i} as ($1/c$ times) the energy flux of gravitational waves, and if we integrate this over a suitable surface containing some system radiating gravitational waves, the resulting "luminosity" will necessarily be equal to the rate of loss of that system's energy, provided that by energy we mean $P^0_{\text{mat}} + P^0_{\text{grav}} = P^0_{\text{tot}}$.