# Gravitational waves on the back of an envelope

Bernard F. Schutz[a)]
*Department of Physics, Washington University, St. Louis, Missouri 63130*

(Received 20 June 1983; accepted for publication 25 July 1983)

Using only Newtonian gravity and a little special relativity we calculate most of the important effects of gravitational radiation, with results very close to the predictions of full general relativity theory. Used with care, this approach gives helpful back-of-the-envelope derivations of important equations and estimates, and it can help to teach gravitational wave phenomena to undergraduates and others not expert in general relativity. We use it to derive the following: the quadrupole approximation for the amplitude $h$ of gravitational waves; a simple upper bound on $h$ in terms of the Newtonian gravitational field of the source; the energy flux in the waves, the luminosity of the source (called the "quadrupole formula"), and the radiation reaction in the source; order-of-magnitude estimates for radiation from supernovae and binary star systems; and the rate of change of the orbital period of the binary pulsar system. Where our simple results differ from those of general relativity we quote the relativistic ones as well. We finish with a derivation of the principles of detecting gravitational waves, and we discuss the principal types of detectors under construction and the major limitations on their sensitivity.

## I. INTRODUCTION

In recent years gravitational radiation has assumed new importance in astronomy, with the observation[1] of its effects in the binary system containing the pulsar PSR1913 + 16 and the inference[2] that the same effects control cataclysmic variables. In the near future this importance will increase as the next generation of ultrasensitive gravitational wave detectors either observes or places useful upper limits on the radiation from supernova explosions. Although gravitational radiation is one of the consequences of general relativity, it is not necessary to understand general relativity in order to understand what this radiation is and to derive order-of-magnitude estimates of its effects. The aim of this paper is to show how to derive these effects from nothing more than a knowledge of Newtonian gravity and a little special relativity,[3] and to develop formulas suitable for "back-of-the-envelope" estimates of the size of important numbers. (These formulas are highlighted by the symbol ♦ on the left-hand margin.) In most cases the results are very close to the predictions of general relativity; this is testimony to the fact that Einstein stuck as close to Newtonian theory as special relativity and the equivalence principle would allow. Where our results differ from Einstein's, I will be careful to point them out and quote the general-relativistic ones.

Most of the discussion should be accessible to intermediate-level undergraduates. For further reading on general relativity itself, see the primer published in this journal by Price[4] or any number of introductory texts. Also in this journal, Campbell and Morgan[5] have described how general-relativistic gravitational radiation may be treated in a manner analogous to the electric/magnetic decomposition of electromagnetic radiation. Relativistic versions of these estimates may be found in the review by Douglass and Braginski.[6]

## II. WHAT ARE GRAVITATIONAL WAVES?

Newtonian gravity is based on the field equation for the Newtonian potential $\phi(\mathbf{x},t)$

$$\nabla^2\phi = 4\pi G\rho, \tag{1}$$

where $\rho$ is the mass density of the source of the field and $G$ is Newton's gravitational constant, $G = 6.673\times 10^{-11}$ m$^3$ kg$^{-1}$ s$^{-2}$. This equation has the solution (which we will call the Newtonian field $\phi_N$)

$$\phi_N(\mathbf{x},t) = -G\int \rho(\mathbf{y},t)r^{-1}\,d^3y,\ r\equiv|\mathbf{x}-\mathbf{y}|, \tag{2}$$

which may be thought of as a superposition of the $1/r$ fields of each element of mass $\rho\,d^3y$ at position $\mathbf{y}$. If the source is changing with time, as in the case of two stars orbiting each other, then Eq. (2) implies that the change in $\phi$ produced by the change in $\rho$ is *instantaneous*: The time $t$ is the same on both sides of the equation. But special relativity tells us that no information should be able to propagate faster than the speed of light $c$. In order to make gravity consistent with this principle, the simplest thing we can do is to modify Eq. (2) to put delay (a retardation) between the time on the right-hand side and that on the left-hand side: A change in $\rho$ at $\mathbf{y}$ ought to be felt at $\mathbf{x}$ only after a time $|\mathbf{x}\text{-}\mathbf{y}|/c$. This leads to the modified field (which we call the relativistic field $\phi_R$)

$$\phi_R(\mathbf{x},t) = -G\int \rho(\mathbf{y},t-r/c)r^{-1}\,d^3y. \tag{3}$$

*We will take this as our fundamental equation and derive all the effects of gravitational radiation from it.* We will discuss later what relation it bears to general relativity, which is a more sophisticated modification of (2) that still embodies the fundamental retardation property of (3). It is easy to show (but is unimportant for our purposes) that $\phi$ in (3) satisfies the scalar wave equation,

$$\nabla^2\phi - \frac{1}{c^2}\frac{\partial^2\phi}{\partial t^2} = 4\pi G\rho.$$

The simple insertion of retardation in (3) is responsible for gravitational waves. To see how this comes about, consider the spatial gradient of $\phi$:

$$\nabla\phi_R = G\int\left(\frac{\rho}{r} - \frac{1}{c}\frac{\partial\rho}{\partial t}\right)\frac{\mathbf{x}-\mathbf{y}}{r^2}\,d^3y. \tag{4}$$

The region of integration is limited to the bounded region in which $\rho\neq 0$, a region of dimension $R$, say. Suppose we let the origin of coordinates be in this region. If we consider $\nabla\phi_R$ far from the source, so that $|\mathbf{x}| \gg R$, then $r \approx |\mathbf{x}|$ and we can take $r$ so large that the first term in Eq. (4) is negligible compared to the second. Then

$$\mathbf{n}\cdot\nabla\phi_R \approx -c^{-1}\partial\phi_R/\partial t, \tag{5}$$

where

$$\mathbf{n} \equiv \mathbf{x}/|\mathbf{x}| \tag{6}$$

is the unit radial vector from the origin (inside the source) to the field point $\mathbf{x}$. Equation (5) says that if we are far away, the typical length scale $\lambda$ on which $\phi_R$ changes ($\mathbf{n}\cdot\nabla\phi_R \sim \phi_R/\lambda$) is $c$ times the typical time scale $P$ on which $\phi_R$ (and the source $\rho$) changes ($\partial\phi_R/\partial t \sim \phi_R/P$). This is characteristic of a wave traveling at speed $c$. In contrast to the Newtonian potential $\phi_N$ in Eq. (2), whose spatial length scale is determined only by $|\mathbf{x}|$ independently of $\partial\rho/\partial t$, the relativistic $\phi_R$ has a spatial dependence which "forgets" the distance $|\mathbf{x}|$ to the source and is sensitive only to $\partial\rho/\partial t$. This is a gravitational wave.

It will turn out that the time-dependent part of $\phi_R$ will often be a small fraction of $\phi_R$. Since $\phi_R$ has the dimensions of (velocity)$^2$, it is conventional to define a dimensionless quantity

$$h \equiv (\text{time-dependent part of } \phi_R)/c^2. \tag{7}$$

Discussions of gravitational-wave detection in the literature commonly refer to $h$, the amplitude of the wave. This is the same $h$ as we have defined in Eq. (7), at least to within a factor of 2.

## III. GENERATION OF GRAVITATIONAL WAVES

### A. The quadrupole approximation

We shall calculate the dominant contribution to $h$ far from the source ($|\mathbf{x}| \gg R$) when the motions in the source are slow compared to $c$. This involves two approximations in Eq. (3):

(i) The overall factor $r^{-1}$ is nearly $|\mathbf{x}|^{-1}$:

$$r^{-1} = |\mathbf{x}|^{-1} + 2\mathbf{y}\cdot\mathbf{n}|\mathbf{x}|^{-2} + \cdots .$$

Since we are interested in $|\mathbf{x}| \gg |\mathbf{y}|$, we can approximate

$$\phi_R = -G|\mathbf{x}|^{-1}\int \rho(\mathbf{y}, t\text{-}r/c)d^3y + O(|\mathbf{x}|^{-2}). \tag{8}$$

(ii) The retarded time $t\text{-}r/c$ is nearly equal to

$$t_0 \equiv t - |\mathbf{x}|/c, \tag{9}$$

and since we assume that $\rho$ changes slowly in time we may expand in the difference

$$\begin{aligned} t - t_0 &= (|\mathbf{x}| - |\mathbf{x}-\mathbf{y}|)/c \\ &= -\mathbf{n}\cdot\mathbf{y}/c + O(|\mathbf{x}|^{-1}). \end{aligned}$$

Then Eq. (8) becomes, with dots denoting time derivatives,

$$\begin{aligned} \phi_R = -G|\mathbf{x}|^{-1}\int [&\rho(t_0) - c^{-1}\dot\rho(t_0)\mathbf{n}\cdot\mathbf{y} \\ &+ \tfrac{1}{2}c^{-2}\ddot\rho(t_0)(\mathbf{n}\cdot\mathbf{y})^2 + \cdots]d^3y + O(|\mathbf{x}|^{-2}). \end{aligned} \tag{10}$$

We have suppressed the explicit dependence of $\rho$ on $\mathbf{y}$. We have kept in (10) only terms that fall off as $|\mathbf{x}|^{-1}$. Let us examine each term in succession.

The first term involves only the mass $M$ of the source

$$\int \rho(t_0)d^3y = M = \text{const.} \tag{11}$$

This gives the Newtonian part of the field. Since $M$ is constant it does not contribute to $h$. The second term can be reduced by using the continuity equation

$$\dot\rho(t,\mathbf{y}) + \frac{\partial}{\partial y_j}(\rho V_j) = 0, \tag{12}$$

where repeated subscripts imply a sum over all three coordinates. The second term involves the integral (recall that $\mathbf{n}$ is independent of $\mathbf{y}$)

$$\begin{aligned} \int \dot\rho y_i\, d^3y &= -\int y_i \frac{\partial}{\partial y_j}(\rho V_j)d^3y \\ &= \int \delta_{ij}\rho V_j\, d^3y \\ &= \int \rho V_i\, d^3y = P_i = \text{const.} \end{aligned} \tag{13}$$

(The second step used Gauss's theorem for a volume enclosing the source, so that $\rho = 0$ on its boundary.) So the second term gives only the momentum of the source, which is again constant and thus not part of $h$.

The third term in Eq. (10) gives our first contribution to $h$, which is why we stopped the expansion there. If we define the *quadrupole tensor* $I_{ij}$ of the source,

$$I_{ij}(t) \equiv \int \rho(y,t)y_i y_j\, d^3y, \tag{14}$$

then in (10) we have

$$\int \ddot\rho y_i y_j\, d^3y = \ddot I_{ij}. \tag{15}$$

Putting all this together gives

$$\begin{aligned} \phi_R &\simeq -\frac{GM}{|\mathbf{x}|} + \frac{Gn_iP_i}{c|\mathbf{x}|} - \frac{G}{2c^2}\frac{\ddot I_{ij}n_in_j}{|\mathbf{x}|}, \\ h &\simeq -(G/2c^4)\ddot I_{ij}n_in_j/|\mathbf{x}|. \end{aligned} \tag{16}$$

This is our basic approximation to $h$. Two points are worth noting because they are exactly the same in general relativity. First, $h$ depends on the second time derivative of the quadrupole tensor. Second, the lower-order terms in the slow-motion expansion (10) were eliminated by the conservation laws for mass and momentum. It is interesting to notice that we could have gone through all of this for electromagnetism, interpreting $\rho$ as the charge density. Then $M$ in Eq. (11) would be $Q$, the total charge, again constant. But Eq. (13) would be the integral of the current density, which is *not* constant, and which is easily seen to be the first time derivative of the dipole moment $d_i = \int \rho y_i\, d^3y$. So the dominant contribution to electromagnetic waves comes from the charge dipole moment. In gravity, conservation of momentum eliminates the analogous term, so the radiation comes from the mass quadrupole moment.[7]

But Eq. (16) differs from the general-relativistic result in an important way. Equation (16) depends only on the component of $I_{ij}$ *along* the vector $\mathbf{n}$, which is the direction of propagation of the wave [see Eq. (5)]. In general relativity the waves depend on the *transverse* components of the *trace-free quadrupole tensor*

$$\text{\it I}\!\!\!-_{ij} = I_{ij} - \tfrac{1}{3}\delta_{ij}I_{kk}, \tag{17}$$

where $\mathrm{I}_{kk}$ is the trace of I. Only the components of $\text{\it I}\!\!\!-_{ij}$ orthogonal to $n_i$ contribute to the physical effects of general-relativistic waves. This is a fundamental distinction. Our model gravitational field $\phi_R$ is a scalar, and therefore has only longitudinal waves. General relativity is a tensor theory and has only transverse waves. (The electromagnetic analogy mentioned in the previous paragraph would also break down on this point: Electromagnetic waves are also transverse and depend only on the component of $d_i$ perpendicular to $n_i$.) The fact that $\mathcal{I}_{ij}$ rather than $I_{ij}$ generates the waves in general relativity also has an important consequence: *Spherically symmetric motions generate no gravitational waves*. This is because the only spherically symmetric tensor is a multiple of the unit tensor $\delta_{ij}$, since it is the only matrix which is left unchanged by every orthogonal transformation of the coordinates (every rotation). But if $I_{ij}$ is proportional to $\delta_{ij}$, then $\mathcal{I}_{ij}$ vanishes and there will be no radiation.

### B. Order-of-magnitude estimates

These reservations regarding our model are unimportant if we only use Eq. (16) for order-of-magnitude estimates of $h$. Manipulations similar to those that led to Eq. (13) give, as a rough estimate,[8]

$$\ddot{I}_{ij} \sim 2 \int \rho V_i V_j \, d^3y.$$

In most sources of waves in astrophysics, the velocities are the result of gravitational forces, e.g., those which cause a star to collapse. By the gravitational virial theorem, then, we should expect that the typical value of $V_i V_j$ should be of order $V^2$, which would be bounded by $\phi_{\text{int}}$, some typical value of $\phi$ inside the source. This leads to

$$|\ddot{I}_{ij} n_i n_j| \lesssim 2M\phi_{\text{int}}, \tag{18}$$

and from (16)

$$\blacklozenge \; |h| \lesssim \phi_N \phi_{\text{int}}/c^4, \tag{19}$$

where $\phi_N$ is the Newtonian potential $-GM/|\mathbf{x}|$ of the source at the observer. This simple formula gives very useful upper limits on $h$, as we will see shortly.

### C. Gravitational-wave luminosity

It is obvious on physical grounds that when a gravitational wave passes through a body, the time-dependent Newtonian gravitational forces can cause internal motions in the body and therefore increase the body's internal energy. (We will see how this happens in detail in Sec. IV.) Where has this energy come from? Ultimately it comes from the source of the waves, but if energy is to be conserved at all times, then it must be present in the waves after they are created by the source and before they transfer it to the distant body. (This discussion is, strictly speaking, inappropriate in general relativity, where detailed conservation of energy sometimes fails. But it is sufficient for motivating the present calculation.) The loss of energy by the source is called radiation reaction and is treated in Sec. III D. Here we consider the radiation flux.

How much energy is carried by the waves? The flux of our scalar field in the direction $\mathbf{n}$ is[9]

$$F = -(1/4\pi G)(\mathbf{n}\cdot\nabla\phi_R)\dot{\phi}_R. \tag{20}$$

In the wave zone [i.e., where Eq. (15) applies] we get

$$F = (1/4\pi Gc)\dot{\phi}_R^2 = (c^3/4\pi G)\dot{h}^2. \tag{21}$$

The luminosity of the source is therefore

$$\blacklozenge \; L = 4\pi|\mathbf{x}|^2 F = (c^3/G)|\mathbf{x}|^2\dot{h}^2 \sim (c^3/G)\omega^2|\mathbf{x}|^2 h^2, \tag{22}$$

where the last step is appropriate for monochromatic waves of frequency $\omega$. This is useful for deducing $h$ if we know $L$. For example, from the spin-down rate of the Crab pulsar PSR0531 + 21, we know that its luminosity in gravitational waves cannot exceed its total energy loss rate of $2\times10^{31}$ J s$^{-1}$. At the Crab's distance of 1 kpc and with a radiation frequency equal to 380 s$^{-1}$ (twice its rotational frequency, since the radiation is quadrupole), Eq. (22) sets an upper limit of

$$h_{\text{Crab}} \lesssim 6\times10^{-25} \tag{23}$$

on the amplitude of radiation from the Crab.

In the spirit of Eq. (19) we can get a back-of-the-envelope upper limit for $L$. For a self-gravitating source, $\omega^2$ is bounded above by the natural dynamical frequency $GM/l^3$, where $l$ is the typical size of the source, so $\omega^2 \lesssim \phi_{\text{int}}^3/(GM)^2$. Putting this into (22) and replacing $|\mathbf{x}|^2/(GM)^2$ by $\phi_N^{-2}$, we find

$$\blacklozenge \; L \lesssim L_0(\phi_{\text{int}}/c^2)^5,$$

$$L_0 \equiv c^5/G = 3.629\times10^{52}\ \text{J s}^{-1}. \tag{24}$$

The luminosity of a source is therefore a very strong function of its internal self-gravity.

We can deduce another important formula from (21) by using Eq. (16):

$$F \sim (G/16\pi c^5)\dddot{I}^2/|\mathbf{x}|^2, \tag{25}$$

where we have suppressed indices on $I$, because, as we remarked earlier, our model picks out the wrong components anyway. Then we find

$$L \sim (G/4c^5)\dddot{I}^2. \tag{26}$$

This is remarkably close to the result for general relativity, which is generally known as the *quadrupole formula*

$$L = (G/5c^5)\dddot{\mathcal{I}}_{jk}\dddot{\mathcal{I}}_{jk}. \tag{27}$$

### D. Radiation reaction

How does the emission of this energy affect the source? To compute these radiation-reaction terms we return to Eq. (3) and compute the field inside the source. Different regions of the source interact with each other with a slight retardation delay. As is characteristic of differential-delay equations, this interaction dissipates energy, precisely the energy lost to radiation. So now in (3) we assume that $|\mathbf{x}|$ is of the same order as $|\mathbf{y}|$, but we continue to assume slow motions and expand $\rho(t-r/c)$ about $\rho(t)$:

$$\phi_R = -G\int r^{-1}\sum_{n=0}^{\infty}\left(-\frac{r}{c}\right)^n\frac{1}{n!}\frac{d^n}{dt^n}\rho(\mathbf{y},t)\,d^3y. \tag{28}$$

This is called a "near-zone" expansion, in contrast to the "far-zone" expansion we used in Eq. (10). We need to keep the first six terms, which we now examine in succession. The first term is, of course, the Newtonian term

$$\phi_N = -G\int \rho r^{-1}\, d^3y. \tag{29}$$

In the second term the factors of $r$ cancel and we have

$$\int \dot{\rho}\, d^3y = 0.$$

The third term will be called the first post-Newtonian term in $\phi_R$,

$$\phi_{\text{pN}} = -\frac{G}{2c^2}\int r\ddot{\rho}\, d^3y. \tag{30}$$

The fourth term ($n = 3$) turns out to be independent of $\mathbf{x}$:

$$\int r^2 \dddot{\rho}\, d^3y = \int (|\mathbf{x}|^2 - 2\mathbf{x}\cdot\mathbf{y} + |\mathbf{y}|^2)\dddot{\rho}\, d^3y$$
$$= |\mathbf{x}|^2 \int \dddot{\rho}\, d^3y - 2\mathbf{x}\cdot \int \mathbf{y}\dddot{\rho}\, d^3y$$
$$+ \int |\mathbf{y}|^2 \dddot{\rho}\, d^3y. \tag{31}$$

The first integral vanishes by mass conservation, the second by momentum conservation. The third term is independent of $\mathbf{x}$. Since gravitational forces come from gradients of $\phi$, we will drop this term from now on. (If we were doing electromagnetism, then as remarked earlier the second term would *not* vanish and this term would give the first radiation-reaction effects, proportional to $\dddot{d}_i$.)

The fifth term is our second post-Newtonian term

$$\phi_{\text{ppN}} = -\frac{G}{24c^4}\frac{d^4}{dt^4}\int \rho r^3\, d^3y. \tag{32}$$

The "post-Newtonian" nomenclature is taken from the analogous near-zone expansion of general relativity.

The sixth term will turn out to give radiation reaction. In relativity it is called the $2\frac{1}{2}$-post-Newtonian term, since in the expansion it is only half as far beyond second post-Newtonian as that term is from the first post-Newtonian term. Similar considerations as we applied to Eq. (31) give

$$\phi_{\text{react}} = (G/30c^5)\,(x_i x_j I^{(5)}_{ij} + \tfrac{1}{2}|\mathbf{x}|^2 I^{(5)}_{kk} - x_i T^{(5)}_i), \tag{33}$$

where

$$T_i = \int \rho\, y_i |\mathbf{y}|^2\, d^3y \tag{34}$$

and where $I^{(5)}_{ij}$ is the fifth time derivative of $I_{ij}$.

To see that this is the reaction term let us calculate the work done by $\phi_R$ on the whole source:

$$\frac{dE}{dt} = -\int \rho V_i \nabla_i \phi_R\, d^3x = -\int \dot{\rho}\phi_R\, d^3x. \tag{35}$$

The Newtonian term contributes

$$-\int \dot{\rho}\phi_{\text{N}}\, d^3x = G\int\int \frac{\dot{\rho}(t,\mathbf{x})\rho(t,\mathbf{y})}{|\mathbf{x}-\mathbf{y}|}\, d^3y\, d^3x$$
$$= \frac{d}{dt}\left(\tfrac{1}{2}G\int\int \rho(t,x)\rho(t,y) r^{-1}\, d^3y\, d^3x\right), \tag{36}$$

which is minus the rate of change of the Newtonian potential energy, as we expect. But this energy is not lost to radiation, as we can see if we consider a *periodic* source and average over one period $P$, defining

$$\langle f\rangle \equiv \frac{1}{P}\int_0^P f\, dt. \tag{37}$$

Then we find

$$\left\langle -\int \dot{\rho}\phi_{\text{N}}\, d^3x\right\rangle = 0, \tag{38}$$

because (36) is a total time derivative integrated between times when the system is in identical configurations. For the next term we find

$$\left\langle -\int \dot{\rho}\phi_{\text{pN}}\, d^3x\right\rangle = \frac{G}{2c^2}\left\langle \int\int r\dot{\rho}(t,\mathbf{x})\ddot{\rho}(t,\mathbf{y}) d^3y\, d^3x\right\rangle$$
$$= \frac{G}{4c^2}\left\langle \frac{d}{dt}\int\int r\dot{\rho}(t,\mathbf{x})\dot{\rho}(t,\mathbf{y}) d^3y\, d^3x\right\rangle$$
$$= 0$$

and similarly

$$\left\langle -\int \dot{\rho}\phi_{\text{ppN}}\, d^3x\right\rangle = 0.$$

But the reaction term is different:

$$\left\langle -\int \dot{\rho}\phi_{\text{react}}\, d^3x\right\rangle$$
$$= -(G/30c^5)\left\langle \dot{I}_{ij} I^{(5)}_{ij} + \tfrac{1}{2}\dot{I}_{jj} I^{(5)}_{kk} - T^{(5)}_i \int \dot{\rho} x_i\, d^3x\right\rangle.$$

The final average vanishes (momentum conservation again) and we finally get, after integrating twice by parts on time,

$$\left\langle \frac{dE}{dt}\right\rangle = -(G/30c^5)\langle \dddot{I}_{ij}\dddot{I}_{ij} + \tfrac{1}{2}\dddot{I}_{jj}\dddot{I}_{kk}\rangle.$$

This is exactly the radiated luminosity that we would have obtained in Eq. (26) had we kept track of our indices, so the energy lost in the source ($\langle dE/dt\rangle$ is negative) appears in waves at infinity. The correct general-relativistic formula is given by (27):

$$\left\langle \frac{dE}{dt}\right\rangle_{\text{gen. rel.}} = -(G/5c^5)\langle \dddot{I\!\!\!-}_{jk}\dddot{I\!\!\!-}_{jk}\rangle. \tag{39}$$

The local reaction potential in general relativity is simpler than (33):

$$\phi_{\text{react, gen. rel.}} = (G/5c^5) x_i x_j I\!\!\!-^{(5)}_{ij}. \tag{40}$$

The reaction force is the negative of the gradient of this potential.

The whole question of the energy radiated by sources in general relativity and of the radiation reaction they experience has been embroiled in controversy for some time, for reasons which would be inappropriate to discuss at length here. The controversy is subsiding, however, and there is now nearly unanimous agreement[10] that Eqs. (27),(39), and (40) can be used as any astrophysicist would want to use them in estimating the strength of waves emitted by a slow-motion source and the effects of that emission on the source.

### E. Sources: supernovae

Let us now put these formulas to work in deriving order-of-magnitude estimates for a few sources. From the point of view of detecting the radiation, the most promising class of sources seems to be *supernova explosions*. These occur when a star uses up its nuclear fuel and is unable to support itself against its own gravity. The central core of the star, with mass about $1M_\odot$, collapses inwards. At great density part of the collapsing matter "bounces," expelling the remaining envelope of the star (which may have a mass of 6–$20M_\odot$). Most of the collapsed matter remains collapsed, as either a black hole or a neutron star. Some nuclear energy is released in the explosion, but most energy comes from the gravitational binding energy of the collapsed remnant, so Eq. (19) applies. An upper limit to $\phi_{\text{int}}$ is its value for a massive neutron star, $\phi_{\text{int}}/c^2 \sim 0.2$. To compute $\phi_{\text{N}}$ we need to decide how far away the supernovae is. A galaxy will have a supernova only once every 20–30 years, so if we wish a reasonable event rate for terrestrial detectors we have to be able to detect distant explosions. The Virgo cluster of galaxies has approximately one event per month. A $1M_\odot$ core at the distance of $d_{10}\times 10M\text{pc}$ ($\sim 3\times 10^{23} d_{10}$ m) has

$$|\phi_N/c^2| = \frac{GM_\odot/c^2}{d_{10}\times 10M\text{pc}} \sim \frac{1.5\times 10^3\ \text{m}}{3\times 10^{23}\ \text{m}} \cdot d_{10}^{-1}$$
$$\sim 5\times 10^{-21} d_{10}^{-1}.$$

Therefore we find

$$\blacklozenge\ |h_{\text{supernova}}| \lesssim 10^{-21} d_{10}^{-1}. \tag{41}$$

This magic figure of $10^{-21}$ is the sensitivity being aimed for in the next generation of detectors (see below). If we are lucky enough to have a supernova in our own galaxy, then this upper limit rises to $\sim 10^{-18}$.

But this could be a serious overestimate, because a perfectly spherical collapse would produce *no* radiation. What effects might prevent a spherical collapse? Rotation is one, and all stellar cores are probably rotating. Tidal interaction between binary stars is another: If an explosion takes place in a close binary, the core may be distorted by a few percent by the gravitational field of the companion. Collapse amplifies such asymmetries, but at present the hydrodynamical calculations are beyond the best computers, so we have little theoretical basis for estimating a more realistic value for $h$. This is one reason for building gravitational-wave detectors: They will put observational constraints on the asymmetry of collapse.

Are there other observational constraints? Katz[11] pointed out a number, one of which provides a *lower bound* on $h$. Pulsars, which are presumably produced in most supernova explosions, have an unusually high spatial velocity,[12] on the order of 120 km s$^{-1}$, compared with tens of km s$^{-1}$ for their progenitors. This must reflect an asymmetry in the collapse that formed them, in which there was a differentiation of momentum within the system, with the neutron star moving one way and the envelope moving the other. Indeed, consider

$$\dot{I}_{ij} = \int \rho(V_i y_j + V_j y_i) d^3y.$$

The moment of momentum implied by the spatial velocity of pulsars means that the right-hand side will not vanish, and so gives a lower bound on $\dot{I}$. We estimate this to be

$$|\dot{I}| \gtrsim (\text{pulsar momentum})\times(\text{size of system})$$
$$\gtrsim m_p V_p l.$$

For $h$ we need $\ddot{I}$, so we assume that this separation of momentum occurred on the collapse time scale $\tau$ (it is hard to imagine it happening on a longer time scale), which leads to

$$|\ddot{I}| \gtrsim m_p V_p l/\tau \sim m_p V_p V_c,$$

where $V_c$ is the collapse velocity. Taking $m_p \sim 1M_\odot$, $V_p \sim 10^2$ km s$^{-1}$, $V_c \sim 0.1c$, we find from (16) for a supernova at a distance $d_{10}\times 10M$ pc

$$\blacklozenge\ |h_{\text{supernova}}| \gtrsim 10^{-25} d_{10}^{-1}. \tag{42}$$

This is a lower limit because other asymmetries can radiate which do not produce a fast-moving remnant. For example, a collapsing core which is rotating but symmetric on reflection through the equatorial plane—as presumably isolated stars are—will not produce a moving remnant because no direction is preferred. At present, it seems that only detections of gravitational waves will tell us where between the two extremes (41) and (42) real collapse falls.

The detectors presently under construction (see below) will not have time resolution shorter than the expected duration of a supernova burst, about 1 ms. From their point of view, therefore, what is of interest is the *total* pulse energy crossing their apparatus in their bandwidth. We shall call this the pulse *strength*. They have invented a unit for this, the *gravitational pulse unit* (GPU):

$$1\ \text{GPU} = 10^2\ \text{J m}^{-2}\ \text{Hz}^{-1}.$$

To relate this to $h$ one needs to make assumptions about the duration and bandwidth of the radiation. For bursts it is reasonable to assume that the bandwidth $B$ is of the same order as the typical frequency $\nu$ and is just the reciprocal of the burst time $\tau$, so that the pulse's strength $S$ is related to its flux $F$ by

$$S = F\tau/B \sim F\tau^2.$$

Using Eq. (21) for $F$ and the two limits (41) and (42) on $h$ gives

$$S \sim \frac{c^3}{4G}\dot{h}^2\tau^2 \sim \frac{c^3}{4\pi G} h^2 \nu^2 \tau^2 \sim \frac{c^3}{4\pi G} h^2,$$

$$\blacklozenge\ 3\times 10^{-18} d_{10}^{-2} \lesssim S\,(\text{GPU}) \lesssim 3\times 10^{-10} d_{10}^{-2}. \tag{43}$$

### F. Sources: binaries

Another much-discussed source of waves is a *binary system*. Here the upper limit (19) is a realistic estimate, since the orbital motion is highly nonspherical. A close binary ($\phi_{\text{int}} \sim 10^{-3}$) in our galaxy ($|\mathbf{x}| \sim 1$ kpc $\sim 3\times 10^{19}$ m) with a total mass of $10M_\odot$ has $|h| \sim 5\times 10^{-19}$. This looks good by comparison with Eq. (41), but we can see from Eq. (22) that the energy flux in the waves depends also on the frequency, and the binary orbital frequencies are in the millihertz region, compared to the kHz frequencies of supernova waves. This low energy flux and technical considerations in the design of detectors make even the closest binaries only marginally detectable.

The *binary pulsar* system[1] contains the pulsar PSR1913 + 16 and another star, presumably also a neutron star, which is unseen. The stars, which each have a mass[13] of $1.4M_\odot$, are separated by a distance of $\sim 1.2\times 10^9$ m. This makes $\phi_{\text{int}} \sim 2\times 10^{-6} c^2$. If the system is 5 kpc $\sim 1.5\times 10^{20}$ m away, then $h \sim 6\times 10^{-23}$, which is far too small for direct detection at present. But the energy lost by the system is significant. Equation (24) now gives a realistic estimate,

$$L_{1913+16} \sim 10^{24}\ \text{J s}^{-1}.$$

A more revealing number is the energy-loss time scale,

$$\tau_E \equiv |E/\dot{E}| \sim \frac{GM}{c^3}(\phi_{\text{int}}/c^2)^{-4}.$$

The number $GM/c^3$ is the light-travel time across the gravitational radius of the system, $GM/c^2$. For the binary pulsar we find $\tau_E \sim 1.3\times 10^{10}$ yr. What is actually measurable is the period change $\dot{P}$. For Newtonian orbits $P \sim |E|^{-3/2}$, so

$$\tau_P \equiv |P/\dot{P}| = \tfrac{2}{3}\tau_E \tag{44}$$

which is then $\sim 10^{10}$ yr for the binary pulsar. A more accurate calculation taking into account the large ellipticity of the orbit, which causes significantly more radiation from the higher velocity at periastron (a correction of a factor of

12 in this case) and using the full general relativistic formulas gives a prediction a factor of 25 smaller than this, $4\times 10^8$ yr. The measurements[1] agree with this prediction to within the observational errors of about 10%.

## IV. DETECTION OF GRAVITATIONAL WAVES

### A. How waves affect matter

To build a detector we must understand how gravitational waves affect matter. In particular, how is it possible to detect $h$, which is much smaller than $\phi_N/c^2$, when we cannot detect the original star itself by measuring $\phi_N$? The answer lies in the nature of gravity. First, the Newtonian acceleration is given by the gradient of $\phi$; the acceleration associated with $\phi_N$ is $\phi_N/|\mathbf{x}|$, while the acceleration associated with a wave of wavelength $\lambda$ is of order $2\pi c^2 h/\lambda$. Since $|\mathbf{x}|\gg\lambda$, this is a big gain for the wave. But this acceleration is not directly measurable in the laboratory, since the whole Earth feels the same acceleration. This is the equivalence principle at work: For the same reason, astronauts do not *feel* any acceleration from the Earth's field as they circle in orbit. What we can measure in the laboratory is only the difference in this acceleration across an experiment. This is called the *tidal force*, since it is the difference in the Moon's gravitational force across the Earth which raises the tides. If the experiment has size $l$, then this tidal acceleration is $l$ times another gradient of $\phi$. So the tidal acceleration due to $h$ is

$$\text{tidal acceleration } \sim c^2 hl(2\pi/\lambda)^2 \sim hl\omega^2, \qquad (45)$$

where $\omega$ is the angular frequency of the wave. This is generally much larger than the corresponding number for the Newtonian field, $\phi_N l/|\mathbf{x}|^2$. In fact, comparing them tells us how far away we must be for waves to dominate the source's field:

$$|\mathbf{x}|_{\text{wave-dominated}} \gtrsim \left(\frac{\lambda}{2\pi}\right)\left(\frac{\phi_N}{c^2 h}\right)^{1/2} \gtrsim \left(\frac{\lambda}{2\pi}\right)\left(\frac{\phi_{\text{int}}}{c^2}\right)^{-1/2}. \qquad (46)$$

A simple way of seeing how the waves affect matter is to consider how two free particles in empty space react to the wave. They experience a relative acceleration given by (45) with

$$h = h_0 e^{i\omega t} \qquad (47)$$

at their position. The equation for the change in their separation $\delta l$ is thus

$$\delta\ddot{l} = \omega^2 l h_0 e^{i\omega t},$$

whose solution is $\delta l = \delta l_0 \exp(i\omega t)$, with

$$|\delta l_0/l| = |h_0|. \qquad (48)$$

This gives a physical interpretation of $h$: It is the *relative strain* induced in a system of free particles by the wave. The goal of detecting $h\sim 10^{-21}$ seems, in this light, formidable indeed!

We have ignored so far the direction in which the particles are displaced from one another by the wave. For our scalar model, this is in the direction of propagation of the wave, since in transverse directions the gradient of $h$ is of order $h/|\mathbf{x}|$ rather than $h/\lambda$. But this is not true of general relativity, which as we have already remarked has transverse waves. In general relativity, distances along the direction of propagation are *unaffected* while distances transverse to the wave are changed as in Eq. (48). The exact manner in which this happens and the nature of the polarization states of the wave are discussed in Ref. 4.

### B. Bar detectors

The oldest and most familiar kind of detector is the *bar detector* pioneered by Weber[14] in the 1960's. Although Weber's early reports of detections have not been confirmed by subsequent observations by other groups,[15] Weber's design has formed the basis of a number of detectors of improved sensitivity. The most sensitive currently operating appears to be the cryogenically cooled one at Stanford, developed by Fairbank[16] and collaborators. At its resonant frequency of 842 Hz it has a limiting $h$ of about $3\times 10^{-18}$ (0.1 GPU).

The principle of the bar detector is to use the gravitational tidal force of the wave to stretch a massive cylinder along its axis, and then to measure the elastic energy of vibration gained by the cylinder. In order to reduce the thermal noise amplitude during a short burst of radiation, bars are constructed to have little damping (high $Q$) in their fundamental frequency of longitudinal oscillation, which should be in the range of frequencies of the incoming wave. This means they usually operate as narrow-band detectors, generally somewhere between 500 and 1500 Hz. The typical frequency for a supernova explosion ought to be on the order of the reciprocal of a few light-travel times across a neutron star, which has a radius of about 20 km. This gives about 1000 Hz. But the signal from a supernova should be broadband, so bar detectors ought to be able to see it with any resonant frequency near 1 kHz. Tuning to this frequency means using a bar of $\sim$1–4 m in length.

Bars can now be isolated from their surroundings well enough to make thermal fluctuations (Brownian motion) the principal source of noise. By cooling the bars to a few degrees kelvin and by improving the resonance properties of the material (a quality factor $Q\sim 10^5$ is common, and $10^9$ may be achievable even in these massive systems), an improvement in sensitivity of about four orders of magnitude has been made in the last decade or so. But another four are necessary, and to achieve that bars must face and overcome a difficult barrier called the *quantum limit*: a burst of short duration $\tau$ will typically deposit an energy

$$E \sim \tfrac{1}{2}M\omega^2\delta l^2 \sim \tfrac{1}{2}M\omega^2 l^2 h^2(\tau\omega/2\pi) \sim 2\times 10^{-32}\ \text{J}$$

in a bar of mass $M\sim 10^3$ kg and length $l\sim 1$ m whose frequency is $\omega\sim 2\pi\times 10^3\ \text{s}^{-1}$, for a burst of amplitude $h\sim 10^{-21}$ and duration $\tau\sim 2\pi/\omega$. But the energy of one phonon of excitation of the bar's vibrational mode is

$$E_{\text{ph}} = \hbar\omega \sim 6\times 10^{-31}\ \text{J}.$$

So the excitation energies expected on classical arguments are small fractions of the phonon energy, and the process of measuring the excitation of the bar must be arranged so as to disturb its quantum state very little; in particular, the measurement should *not* leave the bar in an eigenstate of energy. So-called quantum nondemolition measurements *are* possible, and the theory of how to make them is one of the interesting contributions that gravitational wave research has made to the rest of physics.[17] But these measurement schemes will be very difficult to implement, with the result that the sensitivity of bars may soon be overtaken by that of the laser interferometers.

### C. Laser interferometers

Laser interferometers more nearly resemble the free particles that we discussed earlier than bar detectors do. They consist of three independent masses located at the vertices of an isosceles right triangle. When a wave arrives in the direction of one of the legs, it will change the length of the transverse leg but not the longitudinal one. If each leg is the arm of an interferometer of the Michelson type, this relative change in the length of the arms will produce a shift in the interference fringe.

Although simple in concept, the technical difficulties are formidable. Suppose the arm length is $l \sim 100$ m. Then for $h \sim 10^{-21}$, Eq. (48) implies $\delta l \sim 10^{-19}$ m. This is only $\sim 2 \times 10^{-13}$ of the wavelength $\lambda_\gamma$ of visible light, which is the sort of distance one would need in order to shift the interference pattern from destructive to constructive interference. The fundamental limitation on such a measurement will usually be the number $N$ of photons available, since a distance can be measured to an accuracy $\sim \lambda_\gamma / 4N^{1/2}$ with $N$ photons of wavelength $\lambda_\gamma$. If the measurement time is to be 1 ms (in order to achieve a bandwidth of $10^3$ Hz) then the 100 m interferometer would need a power of $10^{27}$ photons/s, or roughly 100 MW. Since powers $\lesssim 1$ W are available in highly stabilized lasers, the obvious way to improve this is to make the apparatus bigger, because the displacement $\delta l$ increases linearly with $l$ (until one reaches the wavelength of the gravitational wave, about $3 \times 10^5$ m). An interferometer with a physical size of 1 km and 300 reflections down each arm has an effective arm length of 300 km, in which case the power requirement goes down to $\sim 50$ W, which is still large. But by "recycling" the light which is normally thrown away from an interferometer when the interference fringe is read, this kind of power in the arms can be retained.[18] Currently, detectors are being built with physical sizes in the 10–40 m range and effective sizes from 1 to 10 km at the California Institute of Technology, the University of Glasgow, and the Max Planck Institute for Quantum Optics in Garching, near Munich, and others are planned.[18] By the end of 1983 one of the above detectors may be near the sensitivity of the Stanford bar. The next generation of laser interferometers will be on the kilometer scale, taking them into the $10^{-21}$ range.

Interferometers offer a large advantage to the astronomer in that they are broadband detectors with a range of roughly 500–2000 Hz. They do not rely on resonance for their sensitivity. This means that they should (if their sensitivity permits) be able to resolve the spectrum of gravitational-wave bursts, which will be an important diagnostic for inferring the detailed dynamics of the source. However interferometers can also in principle be tuned[18] to resonate at a particular frequency, e.g., to look for waves from a particular pulsar.

### D. Other detectors

Other detectors of various designs have been discussed and built. One of the most notable is the Japanese detector tuned to look for radiation from the Crab pulsar.[19] Searches for ultralong wavelength waves [$\lambda \sim 1$ AU (astronomical unit)], which could result from the formation of the $10^6 M_\odot$ black holes which may be present in most galactic centers, are most efficiently carried out in space. A Jet Propulsion Laboratory group[20] has put upper limits on such radiation by examining transponder communications with the Voyager spacecraft which flew past Jupiter and Saturn. Any wave pulse would cause a characteristic change in the frequencies of signals passing between Earth and the satellites. This is potentially a very sensitive and inexpensive way of searching for gravitational waves, but in order to achieve astrophysically useful sensitivities, transponding to the spacecraft would have to be improved to permit two-frequency communication in both directions (Earth to spacecraft and vice versa). NASA has not yet agreed to do that on its interplanetary missions.

## V. CONCLUSIONS

Like black holes and gravitational lenses, gravitational waves are one of the consequences of general relativity which are playing an increasingly important role in astronomy. Wave phenomena of all kinds in different branches of physics bear many similarities to one another, and we have exploited this by using a simple scalar generalization of Newton's theory to understand the basic physics of gravitational waves and to develop a number of back-of-the-envelope formulas for estimating their effects.

The first laboratory detection of a gravitational wave will of course be an event of fundamental importance in physics, but the discussion in this paper should make it clear that it will really be a verification of the relativistic nature of gravity rather than specifically of general relativity: Practically any relativistic theory of gravity will have waves, since they arise simply from the retardation of the Newtonian gravitational field. In this respect, observations of the binary pulsar have already provided a better quantitative test of general relativity's predictions concerning gravitational waves. In the long run, the main reason for trying to detect gravitational waves is for the astrophysical information they carry.

Gravitational waves have already established their importance in understanding some binary star systems,[2] but their real contribution to astronomy awaits the next generation of detectors, which should either see the waves from supernovae or set useful constraints on supernova models. And further in the future is the opening up of an entirely new radiation spectrum. Given the surprises astronomy has had from its initial observations in the radio, x-ray, and $\gamma$-ray bands of the electromagnetic spectrum, it is not unreasonable to expect that gravitational wave "observatories" will show us still more phenomena we never dreamed existed.

## ACKNOWLEDGMENTS

It is a pleasure to acknowledge useful conversations with Clifford Will and Jeffrey Winicour, and to thank James Hough for pointing out an error in an earlier version.

[a] Permanent address: Department of Applied Mathematics and Astronomy, University College, Cardiff, Wales.


1. J. H. Taylor and J. M. Weisberg, Astrophys. J. **253**, 908 (1982); V. Boriakoff, D. C. Ferguson, M. P. Haugan, Y. Terzian, and S. A. Teukolsky, *ibid.* **261**, L97 (1982).
2. B. Paczynski and R. Sienkiewicz, Astrophys. J. **248**, L27 (1981); **268**, 825 (1983).
3. Similar approaches were described but not carried as far as we do here in J. Winicour, Astrophys. J. **182**, 919 (1973) and W. K. Rose, *Astrophysics* (Holt, Rinehart, and Winston, New York, 1973), p. 276.
4. R. H. Price, Am. J. Phys. **50**, 300 (1982).
5. W. A. Campbell and T. A. Morgan, Am. J. Phys. **44**, 356 (1976); **44**, 1110 (1976).
6. D. H. Douglass and V. B. Braginski, in *General Relativity: An Einstein Centenary Survey*, edited by S. W. Hawking and W. Israel (Cambridge

U. P., Cambridge, 1979), p. 90. For an elementary general relativistic derivation see B. F. Schutz, *A First Course in General Relativity* (Cambridge U.P., Cambridge, 1984).

[7]Other theories of gravity, such as the Brans–Dicke and Rosen theories, introduce other fields which can lead to dipole-type radiation. Observations of the effects of radiation in the binary pulsar system (Sec. III F) give a stringent test of those theories. See C. M. Will, *Theory and Experiment in Gravitational Physics* (Cambridge U. P., Cambridge, 1981).

[8]The correct expression is $2\int(\rho V_i V_j + t_{ij})d^3y$, where $t_{ij}$ is the stress tensor (including external body forces).

[9]Apart from the factor of $G^{-1}$, which ensures the correct units, this is the standard formula for any scalar field. It may be deduced from variational-principle arguments.

[10]M. Walker and C. M. Will, Phys. Rev. Lett. **45**, 1741 (1980); T. Futamase, Phys. Rev. D **28**, 2373 (1983).

[11]J. Katz, Nature **283**, 551 (1980).

[12]A. G. Lyne, in *Pulsars*, edited by W. Sieber and R. Wielebinski, proceedings of International Astronomical Union Symposium 95 (Reidel, Dordrecht, 1981), p. 423.

[13]The measurement of the mass of the stars is an interesting sidelight on the use of general relativity in modeling this system. Ordinary measurements on binaries only give the value of a function of the two masses (the mass function). Measurement in this system of various post-Newtonian effects, such as the perihelion precession, gives other relations which suffice to determine the individual masses and the angle of inclination $i$ of the orbit.

[14]J. Weber, Phys. Rev. Lett. **22**, 1320 (1969); J. Weber, in *General Relativity and Gravitation*, edited by A. Held (Plenum, New York, 1980), Vol. 2, p. 435.

[15]See B. L. Brown, A. P. Mills, and J. A. Tyson, Phys. Rev. D **26**, 1209 (1982); and also Ref. 6 above.

[16]S. P. Boughn, W. M. Fairbank, R. P. Giffard, J. N. Hollenhorst, M. S. McAshan, H. J. Paik, and R. C. Taber, Phys. Rev. Lett. **38**, 454 (1977).

[17]C. M. Caves, K. S. Thorne, R. W. P. Drever, V. D. Sandberg, and M. Zimmerman, Rev. Mod. Phys. **52**, 341 (1980); for the analog in laser interferometers, see C. M. Caves, Phys. Rev. D **23**, 1693 (1981).

[18]See R. W. P. Drever, in *Gravitational Radiation*, edited by T. Piran and N. Deruelle (North-Holland, Amsterdam, 1983). See also H. Billing, K. Maischberger, A. Rudiger, S. Schilling, L. Schnupp, and W. Winkler, J. Phys. E **12**, 1043 (1979).

[19]I. H. Hirakawa, K. Tsubono, and M. Fujimoto, Phys. Rev. D **17**, 1919 (1978).

[20]R. W. Hellings, P. S. Callahan, J. D. Anderson, and A. T. Moffet, Phys. Rev. D **23**, 844 (1981).

# Constraints

N. G. van Kampen
*Institute for Theoretical Physics of the University at Utrecht, Utrecht, The Netherlands*

J. J. Lodder
*FOM-Institute for Plasma Physics "Rijnhuizen," Nieuwegein, The Netherlands*

(Received 25 April 1983; accepted for publication 24 August 1983)

The constraints of classical mechanics are idealizations of stiff elastic forces. They can be derived from such forces in the limit of infinite stiffness, but this derivation presents some difficulties and the limit has to be taken carefully. In statistical mechanics it is no longer possible to idealize stiff potentials by constraints, not even approximately, except under a special and unrealistic restriction on the form of those elastic forces. In quantum mechanics the limit in general does not exist unless the same unrealistic condition is satisfied.

## I. INTRODUCTION

Newton formulated his laws of motion for bodies moving freely in space under the influence of forces acting at a distance, such as apples and planets. In terrestrial circumstances, however, the motion of bodies is more often than not hindered by contact with other bodies. A pendulum is prevented from falling freely by a string or rod that connects it with a fixed fulcrum, a billiard ball rolls on a surface, and a train runs along tracks. The modification of Newton's laws needed for dealing with such constrained motions is provided by d'Alembert's principle and was incorporated in a general mathematical formulation by Lagrange.[1] Constraints became an essential part of classical mechanics and were gradually promoted to the status of axioms by Jacobi[2] and Kirchhoff.[3] Hertz[4] went so far as to consider them as a more basic substitute for forces. Mach[5] did not object. Nowadays they are usually introduced as a matter of course.[6,7]

The constraints are merely the result of the elastic forces exerted by connecting strings or rods, or other devices by which the free motion is hindered. These forces differ from those that Newton had in mind by the fact that they are so strong that they barely allow the body under consideration to deviate from a prescribed path or surface. It ought therefore to be possible to *derive* the d'Alembert–Lagrange equations of motion from the unconstrained Newton equations by adding explicitly very stiff elastic forces as a substitute for the constraints. One then expects[8] that the limit of infinite stiffness will reproduce Lagrange's equations. However, we shall show in Sec. II that this limiting process is by no means trivial, and that additional physical considerations are needed to salvage Lagrange's mathematics—once called "the finest example in all science of the art of getting something out of nothing."[9]

In *statistical mechanics* the situation is even worse. The statistical distribution of a system with constraints (in contact with a heat bath) is in general different from that of the same system in which the constraints are replaced with elastic forces. If a polymer in solution is modeled by a string of beads linked by hard rods the average shape is different from that obtained when the rods are elastic. And this dif-

  