Chapter 2

Forces

We consider here forces acting on a single particle, which may be an idealisation of an extended body. Force is what appears on the right hand side of Newton's second law, but one does use Newton's laws to determine whether a force acts: the force between bodies in any given theory is defined as part of the theory.[1] For example, in Newton's theory of gravity, the force between two particles is defined by the inverse square law (see section 2.4) and in electromagnetism the force on a charged particle is the Lorentz force (section 2.3).

We can distinguish two different sorts of forces: contact forces and non-contact forces. A non-contact force is one that acts at a distance. Examples are gravitational and electromagnetic forces. A non-contact force pervades the whole of space: it exists at every point, whether or not there is a particle at that point to feel it. Often, we will refer to such forces as *force fields*.

A contact force is one that the particle experiences by virtue of being in contact with another body. Examples are friction and normal reaction. In fact, for two bodies, these are the only examples: friction is the component of the force between two bodies that lies in the plane of contact; normal reaction is the component of force in the normal direction. Contact forces also occur when a particle moves through a fluid. Contact forces are caused by interactions between the atoms of the two bodies, so are really just convenient idealisations of non-contact forces.[2]

2.1 Potentials

A force is a vector. In three dimensions, it has three components and is therefore determined by three functions. In some very special, but very important cases, these three functions are related and can be expressed in terms of a single function called a *potential*. Potentials are immensely useful, because they are so much easier both to understand and to calculate.

2.1.1 Potentials in one dimension

We consider a force field $F(x)$ acting on a particle. The work done (WD) by the force in moving the particle from position x by an infinitesimal distance dx is, by definition, given by

$$\text{WD} = F(x)dx \qquad\qquad \text{('force times distance moved by force')}$$

The work done by the force in moving the particle directly[3] from x_0 to x is therefore

$$\int_{x_0}^{x} F(x')dx'.$$

The potential, $\phi(x)$, associated with $F(x)$ is defined by

$$\phi(x) - \phi(x_0) = -\int_{x_0}^{x} F(x')dx'. \tag{2.1}$$

[1] Recall that in section 1.2.2, we took the view that Newton's first law determines whether the frame is inertial, which of course assumes that we know whether a force is acting.

[2] 'Contact' has no meaning at the atomic level: atoms don't touch each other. They interact via non-contact forces such van der Waals forces. These forces are very short range compared with, say, gravitational forces which is the essential difference.

[3] No going backwards and then forwards again.

Thus the potential is a measure of the amount of work done on the particle and hence the ability of the particle itself to do work (i.e. to give back the work done on it[4].) Clearly, the potential is only defined up to an additive constant.

2.1.2 Example: uniform gravitational potential

It is often helpful to think of the familiar example of a uniform gravitational field. For a particle of mass m, the force field has magnitude mg and it acts downwards. The work done by the gravitational field when a particle falls from height z_0 to height z is

$$\int_{z_0}^{z} (-mg)dz = mg(z_0 - z)$$

(the minus sign in the integral arises because z is measured upwards but the force acts downwards). Thus

$$\phi(z) - \phi(z_0) = -mg(z_0 - z)$$

and $\phi(z) = mgz + \text{constant}$.

End of example

2.1.3 Total energy

If we differentiate the equality (2.1) with respect to x, we obtain

$$\frac{d\phi}{dx} = -F(x) \tag{2.2}$$

i.e.

'force equals minus gradient of potential'.

For a particle of mass m moving in a force field $F(x)$ with associated potential $\phi(x)$, we define the total energy E of a particle of mass m moving in the potential ϕ by

$$E = \tfrac{1}{2}m\dot{x}^2 + \phi(x) \tag{2.3}$$

the first term being the kinetic energy and the second being the potential energy.[5]

The total energy is conserved in the motion, i.e. independent of time:

$$\begin{aligned}
\frac{dE}{dt} &= m\dot{x}\ddot{x} + \frac{d\phi}{dt} \\
&= \dot{x}F(x) + \frac{d\phi}{dt} && \text{(using the equation of motion, namely Newton's second law)} \\
&= \dot{x}F(x) + \dot{x}\frac{d\phi}{dx} && \text{(by the chain rule)} \\
&= 0. && \text{(by equation (2.2))}
\end{aligned}$$

Thus the work done by the force contributes to the total energy of the particle, as one might expect. The minus sign in the definition (2.1) means that the potential can be thought of as a form of energy stored in the particle by virtue of its position in the force field which is reduced as the force does work on the particle.

We have shown above that total energy, defined by equation (2.3), is a *conserved quantity* when the force on the particle is derived from a time-independent potential according to (2.2). We will find other conserved quantities (such as momentum and angular momentum). To see if a quantity is conserved, all one has to do is differentiate it with respect to time and use the equations of motion (Newton's second law). Conserved quantities do not necessarily exist in more general situations. For example, there is no conserved quantity that could be interpreted as energy in General Relativity.[6]

A potential can provide an understanding of the motion of a particle without having to solve the equations of motion. This is illustrated in the following example.

[4] Rather loosely speaking: particles are structureless objects and can't really do anything

[5] In some situations, the potential is defined not in terms of the force, but in terms of the force on a particle of unit unit mass in a gravitational field, or the force on a particle of unit charge in an electric field.

[6] Conserved quantities are related (by Noether's theorem) to underlying symmetries of the theory. For Newtonian dynamics, the underlying symmetry is the Galilean group (see section 1.1.8). Energy conservation relates to the time translation, momentum conservation relates to spatial translations, and angular momentum conservation relates to rotations.

2.1.4 Example: particle in cubic potential

A particle of unit mass moves in a one-dimensional potential $\phi(x)$, where

$$\phi(x) = x^3 - 3x\,.$$

The force due to this potential is $-\dfrac{d\phi}{dx}$ ('minus the gradient of the potential'), so the equation of motion of the particle is

$$\frac{d^2x}{dt^2} \equiv \ddot{x} = -\frac{d\phi}{dx} = -3x^2 + 3\,. \tag{2.4}$$

Multiplying by $\dfrac{dx}{dt}$ and integrating with respect to time gives the first integral (the energy integral)

$$\tfrac{1}{2}\dot{x}^2 = -\phi(x) + E$$

where E is a constant of integration (the total energy). This first order differential equation can also be integrated in principle to obtain

$$\int \frac{dx}{\sqrt{2E - 2(x^3 - 3x)}} = t.$$

This is an elliptic integral — it cannot be expressed in terms of elementary functions, though its properties have been well-studied.[7]

A more illuminating approach comes from considering the equation of motion (2.4) to be that of a particle of unit mass rolling[8] under the action of gravity in a landscape the height of which above sea-level (say) is $\phi(x)$, as shown in the sketch. (Actually the height is $\phi(x)/g$ so that the gravitational potential is $g \times \phi(x)/g$; but let's just use units in which $g = 1$ so as not to complicate to picture.) Of course, what the particle does is to move along the x-axis, but because the equation of motion is exactly the same, we can translate the problem to that of the rolling particle.

This approach works even for much more complicated potentials, where the integration approach would be unhelpful, and also for potentials that are functions of two variables.

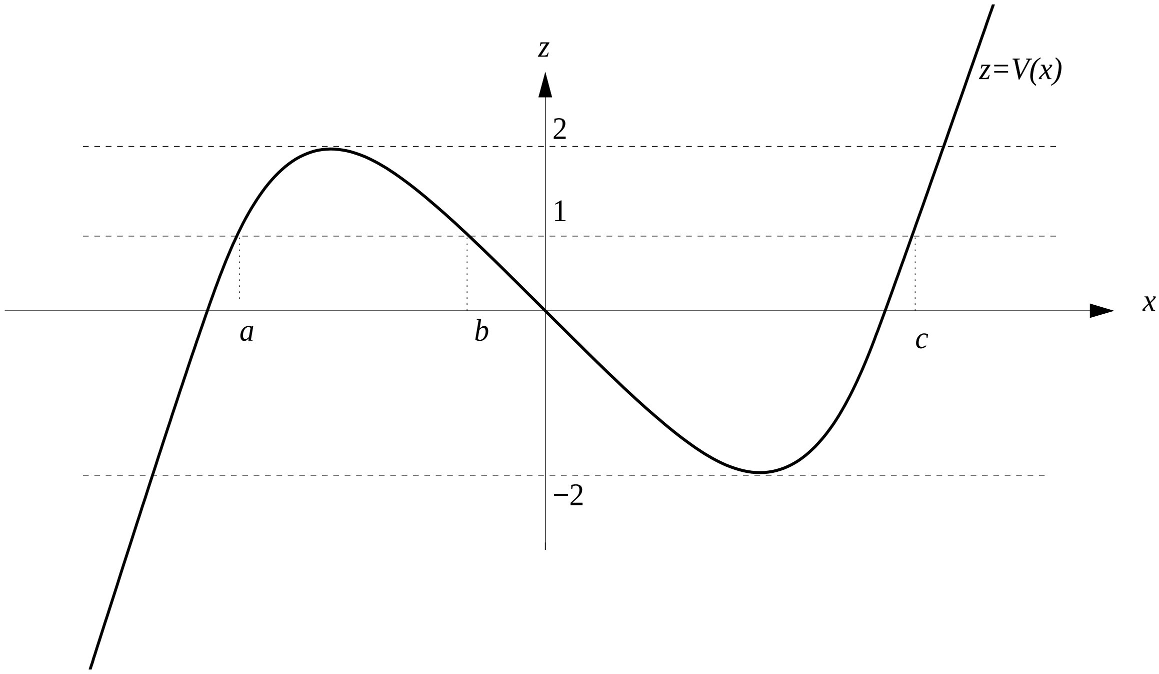
The kinetic energy, and hence speed, of the particle is represented by the difference between the 'height' of the potential function and the fixed 'height' given by the total energy of the particle. At the points where these two heights coincide, the particle has zero speed but non-zero acceleration unless the point is a stationary point of the potential. For a smooth potential function, the particle will reverse when reaching such a point or, if it is a stationary point, will take an infinite amount of time to get there.[9]

---

[7] The result of doing the integral and then expressing x as a function of t gives an *elliptic function*. Elliptic functions are very beautiful mathematical objects, being doubly periodic in the complex plane: they satisfy a relation of the form $F(z+ma+nb) = F(z)$, where z is any complex number (actually, almost any, since elliptic functions generally have singularities), a and b are fixed complex numbers, and m and n are any integers. The functions can be thought of as existing on lattices in the complex plane, or on toruses. The closest familiar analogy are the trigonometric functions, which which are only singly periodic but can also be defined as the inverse of an integral similar to ours (though quadratic rather than cubic or quartic in the square root in the denominator).

[8] Actually, sliding since particles have no size and therefore cannot really be said to roll; but it is normal to call it rolling.

[9] You can easily check this assertion. For a smooth potential, by which I mean a potential with a Taylor series about each point, the motion of the particle very close to the stationary point is determined approximately by the first non-zero term of the Taylor series, i.e. by the equation $\dot{x}^2 = x^n$, where n is an integer greater than 1. Integrating this shows that the time taken to reach $x = 0$ is infinite.



From the diagram, we can see the following possibilities (there are many others), depending on the initial conditions. For convenience, the initial conditions are given in terms of x_0 and E, rather than x_0 and $\dot{x}_0$.

(i) $x_0 < a$, $\dot{x}_0 > 0$, $E = 1$.
In this case, the particle slows down until its velocity is reversed when $x = a$ (see diagram); it then goes off to $x = -\infty$.

(ii) $x_0 = a$, $E = 1$.
The particle, initially stationary, sets off towards $-\infty$, gathering speed.

(iii) $a < x_0 < b$, $E = 1$.
This is not possible: the particle does not have sufficient energy (classically) to exist in this part of the x-axis.

(iv) $b \le x_0 \le c$, $E = 1$.
The particle oscillates between b and c.

(v) $x_0 > c$, $E = 1$.
Again, not possible.

(vi) $E = 3$.
The particle ends up at $-\infty$ either directly if $\dot{x}_0 \le 0$, or after bouncing off the potential if $\dot{x}_0 > 0$.

(vii) $E = 2$, $x_0 = -1$. Note that the turning points of $\phi(x)$ are at ± 1. In this case the particle has no kinetic energy and just stays put. It is in unstable equilibrium, as is obvious from the diagram. This can be checked analytically. Let $x = -1+\epsilon$, where $\epsilon \ll 1$. Then, substituting into the equation of motion (2.4), we have

$$\frac{d^2}{dt^2}(-1+\epsilon) = -3(-1+\epsilon)^2 + 3 \approx +6\epsilon$$

so $\epsilon \approx \epsilon_0 \cosh \sqrt{6}(t - t_0)$, which grows grows exponentially. Small perturbations from the equilibrium will therefore in general become large, which means the equilibrium is unstable.

End of example

2.1.5 Potentials in three dimensions

As mentioned before, we cannot in general expect to be able to express all three components of a force $\mathbf{F}(\mathbf{r})$ in terms of a single potential. An obvious exception is a three-dimensional force that is essentially one-dimensional, such as the uniform gravitational field discussed in a previous example. There are in fact other exceptions, including many of the forces that arise in theoretical physics.

Following the treatment of the one-dimensional case we write the work done by the force in moving a particle from $\mathbf{r}$ to $\mathbf{r} + d\mathbf{r}$ as

$$\mathrm{WD} = \mathbf{F}.d\mathbf{r}.$$

The scalar product arises naturally here because we are only interested in the component of the force in the direction of motion. The work done in moving the particle from $\mathbf{r}_0 \to \mathbf{r}$ is

$$\int_{\mathbf{r_0}}^{\mathbf{r}} \mathbf{F}(\mathbf{r}').d\mathbf{r}' = \int_{t_0}^{t} \mathbf{F}(\mathbf{r}(t')).\frac{d\mathbf{r}}{dt'}dt' \tag{2.5}$$

where, in the second integral, t is a parameter (which could be time) along the path of integration. This is a line integral, and in general its value depends on the path joining $\mathbf{r_0}$ to $\mathbf{r}$. While the integral makes perfect sense as a measure of work done, it does not define a potential function of $\mathbf{r}$, because of the path dependence.

However, for some forces, the value of the integral does not in fact depend on the path. Such forces are said to be *conservative*. For conservative forces, we can define a potential $\phi(\mathbf{r})$ by[10]

$$\phi(\mathbf{r}) - \phi(\mathbf{r_0}) = -\int_{\mathbf{r_0}}^{\mathbf{r}} \mathbf{F}(\mathbf{r}').d\mathbf{r}'.$$

If the path is parameterised by t, we can differentiate this with respect to t, using the second from of the integral in (2.5), to obtain

$$\frac{d\phi}{dt} = -\mathbf{F}(\mathbf{r}).\frac{d\mathbf{r}}{dt} \tag{2.6}$$

By the chain rule,

$$\begin{aligned} \frac{d\phi}{dt} &= \frac{\partial\phi}{\partial x_i}\frac{dx_i}{dt} && \text{(summation convention)} \\ &= \boldsymbol{\nabla}\phi.\frac{d\mathbf{r}}{dt} \end{aligned}$$

Comparing this with (2.6) and remembering that $\frac{d\mathbf{r}}{dt}$, which is the tangent vector to the path, is arbitrary because the path is arbitrary (the value of the integral is the same for all paths), we have

$$\boldsymbol{\nabla}\phi = -\mathbf{F}(\mathbf{r}) \qquad \text{('force equals minus gradient of potential')}$$

As in the one-dimensional case, we define the total energy E of a particle of mass m moving in the potential ϕ by

$$E = \tfrac{1}{2}m\dot{\mathbf{r}}.\dot{\mathbf{r}} + \phi \tag{2.7}$$

and again this is conserved:

$$\begin{aligned} \frac{dE}{dt} &= m\dot{\mathbf{r}}.\ddot{\mathbf{r}} + \frac{d\phi}{dt} \\ &= \dot{\mathbf{r}}.\mathbf{F} + \frac{d\phi}{dt} && \text{(Newton's second law)} \\ &= \dot{\mathbf{r}}.\mathbf{F} + \boldsymbol{\nabla}\phi.\dot{\mathbf{r}} && \text{(chain rule)} \\ &= 0 && \text{(using } \mathbf{F} = -\boldsymbol{\nabla}\phi) \end{aligned}$$

When the potential occurs in the definition of total energy, as in equation (2.7), it is called the *potential energy* of the particle.

2.1.6 Central forces

A force field $\mathbf{F}(\mathbf{r})$ is said to be *central* if it depends only on the distance between the point at which the force is acting (call it $\mathbf{r}$) and a fixed point. If we take the fixed point to be the origin of coordinate, we can write a central force in the form $\mathbf{F}(r)$, where r is the usual spherical polar coordinate.

[10] The definition of a conservative force is one for which the work done by the force is independent of the path for all paths between any two fixed points. We should specify that the fixed points and the paths must lie in some given volume V, which might be the whole of $\mathbb{R}^3$. A consequence (see below) is that there exists a function $\phi(\mathbf{x})$ such that $\mathbf{F} = -\boldsymbol{\nabla}\phi$ in V, and it is easily seen that, if such a function exists, $\mathbf{F}$ is conservative. As will be shown in the Vector Calculus course a necessary and sufficient condition for $\mathbf{F}$ to be conservative is $\boldsymbol{\nabla} \times \mathbf{F} = 0$ in V (i.e. $\mathbf{F}$ is curl-free, or irrotational).

A stronger definition of a central force, which is the one we will adopt, is that it acts towards or away from the fixed point. In this case, we can write the force in the form $f(r)\widehat{\mathbf{r}}$, where $\widehat{\mathbf{r}}$ is the unit vector in the radial direction.

For such a force, we can have hopes that it is conservative, since it depends on one function only; and our hopes are fulfilled. Recall[11] that

$$\boldsymbol{\nabla} r = \widehat{\mathbf{r}}.$$

Thus if we define $\phi(r)$ (up to an additive constant of integration) by $f(r) = -\dfrac{d\phi}{dr}$, we have

$$\mathbf{F}(\mathbf{r}) = f(r)\widehat{\mathbf{r}} = -\frac{d\phi}{dr}\boldsymbol{\nabla} r = -\boldsymbol{\nabla}\phi$$

using the chain rule for the last equality.

Therefore, any central force can be written in terms of a potential, and the total energy (potential plus kinetic) is conserved.

2.2 Friction

As mentioned earlier, friction is a contact force. It is a convenient way of describing the complicated interactions between the atoms of different bodies but is not itself a fundamental force[12]

There are two sorts of friction: *dry friction* and wet friction or *drag*. Dry friction occurs when two bodies are in contact; a particle resting or sliding on an inclined plane, for example. The governing equation is

$$F = \mu R,$$

where R is the normal reaction and μ is the coefficient of friction. This applies both in static friction (a body at rest) and sliding friction, though the coefficient of friction between two given bodies will be different in the two cases. This sort of friction is not particularly relevant to this course.

2.2.1 Fluid drag

Drag occurs when a body is moving through a fluid. Drag is velocity dependent, being normally either linear or quadratic in speed and parallel to velocity.

Linear drag is caused by the stickiness of the fluid and takes the form

$$\mathbf{F} = -k\mathbf{v}$$

where k is independent of velocity. *Stokes's law* for a spherical body gives $k = 6\pi\eta R$, where η is the viscosity and R is the radius of the body. Drag is approximately linear when viscous forces predominate, for example a rock in lava or a bacterium in water.

Quadratic drag takes the form

$$\mathbf{F} = -k|\mathbf{v}|\mathbf{v} \tag{2.8}$$

where k is now depends on the density of the fluid and the cross-sectional area of the body. This occurs when the resistance to motion is due to the body having to push the fluid aside, for example projectiles in air[13] and submarines.

Linear friction dominates when the speed of the body is small in the following sense:

$$\frac{\rho|\mathbf{v}|R^2}{\eta} \approx 1.$$

The dimensionless quantity on the left hand side is called the *Reynolds number*.[14]

---

[11] This is most easily verified using the expression for $\boldsymbol{\nabla}$ in polar coordinates:

$$\boldsymbol{\nabla} f = \frac{\partial f}{\partial r}\widehat{\mathbf{r}} + \frac{1}{r}\frac{\partial f}{\partial \theta}\widehat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \phi}\widehat{\boldsymbol{\phi}}.$$

Otherwise, one can conveniently consider $\boldsymbol{\nabla} r^2$ in Cartesians.

[12] The fundamental forces are gravitation, electromagnetic forces, and weak and strong nuclear forces.

[13] Including stones being dropped from leaning towers — see section 4.

[14] The Reynolds number measures the relative importance of inertial forces and viscous forces in fluid flow: if it is less than about 2000, the flow is laminar; if it is greater than about 4000 the flow is turbulent.

2.2.2 Example: vertical motion under gravity with quadratic friction

A particle of mass m moves vertically under the influence of gravity (assumed uniform) and a quadratic resistance force of magnitude mkv^2, where v is the velocity[15] of the particle. In what follows, we take z to measure distance vertically upwards. By definition, $v = \dfrac{dz}{dt}$, so $v > 0$ if the particle is moving upwards and $v < 0$ if the particle is moving downwards, and similarly $\dfrac{d^2z}{dt^2} > 0$ if v is (strictly) increasing.

(i) Upwards motion

The equation of motion is 'vector' equation — we have to worry about direction just as if we were working in three dimensions. Gravity acts downwards and, when the particle moves upwards, the resistive force also acts downwards. The equation of motion is therefore

$$m\frac{dv}{dt} = -mg - mkv^2, \qquad \text{i.e.} \qquad \frac{dv}{dt} = -g - kv^2.$$

Integrating this will give us v as a function of t. Alternatively, we could write the equation of motion, using the chain rule, as

$$v\frac{dv}{dz} = -g - kv^2 \qquad \text{i.e.} \qquad \tfrac{1}{2}\frac{d(v^2)}{dz} = -g - k(v^2) \tag{2.9}$$

which will give us v^2 as a function of height z.

Suppose the particle is projected upwards from $z = 0$ with speed V and we want to find the maximum height H. We can obtain the form of the expression for H by considering dimensions. The dimension of k is L^{-1}, as can be seen from the equation of motion (2.9). The other relevant quantities are g and V. Since that makes three in total, and they all involve only two dimensions, L and T, there is one dimensionless parameter, call it η. The choice of η is not, of course, unique; one possibility is $\eta = kV^2/g$ (and any function of this quantity would do). We therefore expect that H can be written in the form

$$H = k^{-1}f(\eta),$$

where the function $f(\eta)$ cannot be determined by dimensional analysis.

Integrating equation (2.9), and noting that the particle reaches its maximum height when $v = 0$, gives

$$\int_V^0 \frac{d(v^2)}{g + kv^2} = \int_0^H -2dz$$

so

$$2H = -\frac{1}{k}\log\left(\frac{g}{kV^2 + g}\right) = \frac{1}{k}\log\left(1 + \eta\right). \tag{2.10}$$

Now suppose that effect of friction is weak compared with the effect of gravity. Since the effect of friction is greatest at the point of projection, weak friction corresponds to $kV^2 \ll g$, i.e. $\eta \ll 1$. Expanding the log in the expression for H (2.10) gives

$$H = \frac{1}{2k}(\eta - \tfrac{1}{2}\eta^2 + \cdots) = H_o(1 - \tfrac{1}{2}\eta + \cdots),$$

where H_o is the height that the particle would have attained in the absence of friction. Note that we have to expand to second order in the small parameter to see the affect of friction.

(ii) Downwards motion – e.g. a raindrop

This time the equation of motion is of the particle

$$\frac{dv}{dt} = -g + kv^2,$$

because now the resistive force acts upwards.

Suppose the particle is dropped from (i.e. released from rest at) a great height. As we know from experience, there is a terminal speed which cannot be exceeded; in fact, as we shall see, it cannot be attained.

[15]Note: velocity not speed — it can be either positive or negative

We can find the speed v at any time t by integrating directly[16]:

$$\int_0^t dt = -\int_0^v \frac{dv}{g - kv^2}.$$

Setting $\sqrt{k}\,v = \sqrt{g}\tanh\theta$ gives

$$t = -\frac{1}{\sqrt{gk}}\tanh^{-1}(\sqrt{k/g}\,v) = \frac{1}{\sqrt{gk}}\tanh^{-1}(\sqrt{k/g}\,|v|)$$

(remember that v is negative for a falling particle).

It is a good idea to pause occasionally and check that all is dimensionally in order. Since $[k] = L^{-1}$ and $[g] = LT^{-2}$, we find that $[\sqrt{gk}] = T^{-1}$ and $[\sqrt{k/g}] = L^{-1}T$, which means that the above equation is dimensionally correct.

Thus

$$v = -\sqrt{g/k}\tanh(\sqrt{gk}t).$$

Note that as $t \to \infty$, $v \to \sqrt{g/k}$, though this speed is never attained. The quantity $\sqrt{g/k}$ is called the *terminal velocity* (more properly, the terminal speed).

The quantity $1/\sqrt{gk}$ is the only combination of parameters that has the dimension of time, so it must provide a timescale analogous to the half-life of exponential decay. My calculator give $\tanh 1 \approx 0.76$, so $1/\sqrt{gk}$ is the time taken for the particle to reach about $3/4$ of the terminal velocity, starting from rest.

The terminal velocity for a sky diver in the free fall position (limbs outstretched) is about 55 metres per second — call it 50 metres per second. Taking $g = 10$ metres per second per second, we see that $k^{-1} = 250$ metres and the timescale is 5 seconds. Very roughly, the terminal velocity is proportion to the square root of its area (see the remark following equation (2.8) regarding the dependence of k on area). The terminal velocity for a mouse[17] is much smaller than for a human, so it is more likely to have a happy landing.[18]

End of example

2.2.3 Example: projectile with linear drag, using vectors

As mentioned above, typical projectiles in air are subject to quadratic drag, so the one we are thinking about here must be in water or maybe even treacle.

A particle of mass m is projected from the origin at velocity $\mathbf{u}$. The gravitational acceleration is denoted by $\mathbf{g}$ and the drag force is $-mk\mathbf{v}$, where k is a constant (the m is included here for convenience).

The equation of motion (Newton's second law) is

$$m\frac{d\mathbf{v}}{dt} = m\mathbf{g} - mk\mathbf{v}$$

i.e.

$$\frac{d\mathbf{v}}{dt} + k\mathbf{v} = \mathbf{g}.$$

[16] Note the lazy convention of not using a dummy variable in the integral; the only excuse for this is that there is no scope for confusion here.

[17] J.B.S. Haldane in his essay *On Being the Right Size* summarises the situation nicely but not delicately: 'You can drop a mouse down a thousand-yard mine shaft; and, on arriving at the bottom, it gets a slight shock and walks away, provided that the ground is fairly soft. A rat is killed, a man is broken, a horse splashes.'

[18] Curiously, cats survive big drops better than small drops according to a 1987 study from the Journal of the American Veterinary Medical Association wherein two vets examined 132 cases of cats that had fallen out of high-rise windows and were brought to the Animal Medical Center, a New York veterinary hospital, for treatment.

The vets postulated that cats sense acceleration, rather than speed. When a cat starts falling it begins accelerating at something close to g and it accordingly assumes its 'panic' posture: head tucked in; paws under body, arched back. This protects its vital organs, but unfortunately makes it more aerodynamic i.e. smaller k, corresponding to a bigger terminal velocity at which it is likely to be killed if it strikes a hard surface.

However, the acceleration reduces considerably as terminal velocity is approached, and the cat adopts a different strategy. This tends to happen after the cat has fallen about 8 storeys. It stretches out its legs and neck, like a flying fox, increasing its surface area, which increases k, decreases the terminal velocity, and so slows the cat down. At this lower terminal velocity it can survive the fall — from any height greater than 8 storeys, though apparently 32 stories is the highest on record. No cats were harmed in the making of this footnote.

We can solve this equation using an integrating factor, as if it were an ordinary (non-vector) differential equation. We first rewrite it as

$$\frac{d}{dt}\left(e^{kt}\mathbf{v}\right) = e^{kt}\mathbf{g}$$

then integrate and multiply by e^{-kt}:

$$\mathbf{v} = \frac{1}{k}\mathbf{g} + \mathbf{C}\,e^{-kt}$$

where $\mathbf{C}$ is a constant (vector) of integration which can be identified using the initial condition $\mathbf{v} = \mathbf{u}$ at $t = 0$. Thus

$$\mathbf{v} = \frac{1}{k}\mathbf{g} + (\mathbf{u} - \frac{1}{k}\mathbf{g})e^{-kt}.$$

This equation can be integrated directly to give $\mathbf{r}$:

$$\mathbf{r} = \frac{1}{k}\mathbf{g}t - \frac{1}{k}(\mathbf{u} - \frac{1}{k}\mathbf{g})e^{-kt} + \mathbf{d}$$

where $\mathbf{d}$ is a (vector) constant of integration which can be identified using the initial condition $\mathbf{r} = 0$ at $t = 0$. Thus

$$\mathbf{r} = \frac{t}{k}\mathbf{g} + \frac{1}{k}(\mathbf{u} - \frac{1}{k}\mathbf{g})(1 - e^{-kt}). \tag{2.11}$$

This is the complete solution. Choosing axes such that

$$\mathbf{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \qquad \mathbf{g} = \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix} \qquad \text{and} \qquad \mathbf{u} = \begin{pmatrix} u\cos\alpha \\ 0 \\ u\sin\alpha \end{pmatrix},$$

the solution is

$$x = \frac{1}{k}u\cos\alpha\left(1 - e^{-kt}\right), \quad y = 0, \quad z = -\frac{gt}{k} + \frac{1}{k}\left(u\sin\alpha + \frac{g}{k}\right)\left(1 - e^{-kt}\right).$$

This looks a bit more complicated than the $k = 0$ case, but it is has some expected features. For very large t, in the sense $kt \gg 1$, the exponential terms can be ignored and the particle drops vertically at its terminal speed of g/k. The horizontal component of velocity has been completely eroded by the drag force.

For small k (i.e. $kt \ll 1$) we should retrieve the projectile-without-drag solution. At first sight, this limit looks bad because of the k in the denominator. However, if we expand the exponential in the vector form of the solution (2.11) as far as the quadratic terms we see that the limit is in fact defined (as it must be):

$$\begin{aligned}\mathbf{r} &= \frac{t}{k}\mathbf{g} + \frac{1}{k}(\mathbf{u} - \frac{1}{k}\mathbf{g})(1 - 1 + kt - \tfrac{1}{2}(kt^2) + \cdots) \\ &= \mathbf{u}t + \tfrac{1}{2}\mathbf{g}t^2 + O(kt).\end{aligned}$$

This is of course[19] the solution that we would have obtained by solving the equations of motion with $k = 0$.

End of example

2.3 Motion in an electromagnetic field

2.3.1 The Lorentz force

The *Lorentz force*[20] is the force experienced by a charged particle in an electromagnetic field. It is given by

$$\mathbf{F} = e\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \tag{2.12}$$

[19] It is not quite 'of course'. We are asking if two processes commute: is solving an equation and taking a limiting value of the solution the same as taking the limit in the equation then solving it. This is the sort of thing one has to worry about when studying the theory of differential equations, and partial differential equations in particular, but we needn't let it detain us here.

[20] It was introduced by the Dutch physicist Hendrik Lorentz (after whom the Lorentz transformations of special relativity are named) in 1892, though a more or less equivalent equation occurs in the works of Maxwell thirty years earlier. Lorentz was awarded the Nobel prize in 1902 for his work on the Zeeman effect.

where e is the charge on the particle, $\mathbf{v}$ is the velocity of the particle, $\mathbf{E}$ is the electric field and $\mathbf{B}$ is the magnetic field. The term $e\mathbf{E}$ is called the electric force and the term $e\mathbf{v}\times\mathbf{B}$ is called the magnetic force.[21]

Generally, $\mathbf{E}$ and $\mathbf{B}$ depend on both position and time, though in the simple examples considered here, both these forces are likely to be uniform and constant. Note that the force is defined everywhere in space and time, regardless of whether a charge is present to experience the force.

A positively charged particle will be accelerated in the same direction as the electric field, but will curve perpendicularly to both the instantaneous velocity and the magnetic field according to the right-hand rule.

If we take the scalar product of equation (??) with $\mathbf{v}$, we see that

$$\mathbf{F}\cdot\mathbf{v} = e\mathbf{E}\cdot\mathbf{v}.$$

The left hand side of this equation is the rate at which work is done by the force $\mathbf{F}$ on the particle, and so we see that the magnetic field does not contribute at all the work done; it is all done by the electric part of the force field.

Like drag force, the electromagnetic force $\mathbf{F}$ depends on the velocity of the particle (explicitly), as well as on its position (implicitly, via the dependence of $\mathbf{E}$ and $\mathbf{B}$ on position) so it is not in general conservative; though it may be conservative in special cases[22]

2.3.2 Electric field of a point charge

The electric field of a particle of stationary charge q is given by

$$\mathbf{E} = \frac{q}{4\pi\epsilon_0}\frac{\mathbf{r}}{r^3}. \tag{2.13}$$

where ϵ_0 is a constant called the permittivity of free space. It relates the units of electric charge to the mechanical quantities M, L and T. Its value is $8.8541878210^{-12}\mathrm{m}^{-3}\mathrm{kg}^{-1}\mathrm{s}^4 A$, where A is the basic electric unit (amperes). Since $\boldsymbol{\nabla} r^n = n\mathbf{r}r^{n-2}$ (see the Vector Calculus course), we can write $\mathbf{E}$ as a gradient:

$$\mathbf{E} = -\boldsymbol{\nabla}\left(\frac{q}{4\pi\epsilon_0 r}\right).$$

The quantity $\dfrac{q}{4\pi\epsilon_0 r}$ is called the electrostatic potential for the point charge.

When $\mathbf{B}=\mathbf{0}$, as in the case for a stationary charge, the Lorentz force (2.12) is proportional to $\mathbf{E}$, so the force on a particle moving in the field of a point electric charge is conservative. The field of a point charge is very similar (identical really) to that of a point gravitational mass, as we shall see in the next section.

It is worth noting that since there are no free point magnetic charges (magnetic charges occur in pairs as in a bar magnet), there is no corresponding field for a point magnetic charge. The simplest magnetic field (that is not constant) is called a *dipole* field, which is the result of superposing a positive magnetic charge and a negative magnetic charge.

2.3.3 General motion of a charged particle in an electromagnetic field

In general, $\mathbf{E}$ and $\mathbf{B}$ are functions of both time t and position $\mathbf{r}$; in Cartesian coordinates, they are functions of x_i, where $\mathbf{r}=(x_1,x_2,x_3)$, and t. We assume that the electromagnetic fields are given and are not affected by the presence of a charged particle. [23]

Writing the trajectory of the particle as $\mathbf{r}(t)$, the equation of motion becomes

$$m\ddot{\mathbf{r}}(t) = \mathbf{F} = e\Big(\mathbf{E}\big(\mathbf{r}(t),t\big) + \dot{\mathbf{r}}(t)\times\mathbf{B}\big(\mathbf{r}(t),t\big)\Big),$$

which represents three coupled second-order ordinary non-linear (in general) differential equations with three dependent variables, and can in principle be solved, given suitable initial conditions.

[21] You don't have to know anything about electric and magnetic fields for this course: that comes in Part IB. The Lorentz force is included in this course as an exercise — an important exercise — in handling vector equations of motion.

[22] Electromagnetic fields are governed by the Maxwell equations, one of which is $\boldsymbol{\nabla}\times\mathbf{E} = -\frac{\partial\mathbf{B}}{\partial t}$. Thus the Lorentz force is conservative if (and only if) $\mathbf{B}=\mathbf{0}$.

[23] Of course, this is an idealisation: a moving charged particle will create its own electromagnetic fields and these might well affect the source (whatever it is) of the given $\mathbf{E}$ and $\mathbf{B}$.

2.3.4 Example: motion in a uniform electromagnetic field

Here we consider the case when the electromagnetic field is both constant (in time) and uniform (same at all points in space), so that

$$\frac{\partial \mathbf{E}}{\partial t} = \frac{\partial \mathbf{B}}{\partial t} = \mathbf{0}\,; \qquad \frac{\partial \mathbf{E}}{\partial x_i} = \frac{\partial \mathbf{B}}{\partial x_i} = \mathbf{0}\,.$$

Recall that a vector is (e.g.) time-independent if and only if its Cartesian components are time-independent.

The equation

$$\mathbf{F} \equiv e(\mathbf{E} + \dot{\mathbf{r}} \times \mathbf{B}) = m\ddot{\mathbf{r}}. \tag{2.14}$$

can be tackled in a number of ways. Below, we will solve it entirely in components and also entirely in vectors. Neither method is optimal: a judicious mixture would serve us better.

(i) Component method

The practical way to integrate the questions is to work in components; **BUT** it is essential to choose sensible axes. Since the lines of $\mathbf{B}$ are everywhere parallel, we can choose axes such that the z-axis is parallel to $\mathbf{B}$:

$$\mathbf{B} = (0, 0, B)$$

If $\mathbf{E}.\mathbf{B} = \mathbf{0}$, we can choose axes such that $\mathbf{E} = (E, 0, 0)$, but in general the best we can do (by rotating the x and y axes, which is the only freedom left after fixing the z axis) is

$$\mathbf{E} = (E_1, 0, E_3).$$

With this choice, the equations of motion (2.14) become

$$m\ddot{x} = eE_1 + eB\dot{y} \tag{2.15}$$

$$m\ddot{y} = \qquad - eB\dot{x} \tag{2.16}$$

$$m\ddot{z} = eE_3 \tag{2.17}$$

which can be solved by elementary means or by using matrices.

The solution to third equation (2.17) can be written down:

$$z = (e/2m)E_3t^2 + at + b \tag{2.18}$$

where a and b are constants obtainable from initial conditions.

A neat way to solve the first two equations (2.15) and (2.16), which happens to work in this case, is to set $\xi = x + iy$, and add i times equation (2.16) to equation (2.15); of course, one could always do this to obtain a single complex equation containing both ξ and $\bar{\xi}$, but the special feature of our equations is that the result does not contain $\bar{\xi}$:

$$m\ddot{\xi} = eE_1 - ieB\dot{\xi}\,.$$

This can be integrated straight away:

$$\xi = pe^{-i\omega t} - iE_1t/B + q$$

where $\omega = eB/m$ and the complex constants p and q can be obtained from the initial conditions.[24]

If the particle is initially at the origin, and moving in the y-direction, we find

$$\xi = p(e^{-i\omega t} - 1) - ikt\,,$$

where $k = E_1/B$ and p is real, so

$$x = p(\cos\omega t - 1)\,, \qquad y = -p\sin\omega t - kt\,.$$

This is roughly (exactly if $k = p$) a cycloid, so the motion of the particle is, somewhat counter-intuitively, a uniform acceleration parallel to $\mathbf{B}$ (but due to the component of the electric field parallel to $\mathbf{B}$) and cycloidal motion in the plane perpendicular to $\mathbf{B}$.

[24] ω is called the *Larmor frequency* after the physicist Joseph Larmor, senior wrangler in 1880, Lucasian Professor from 1903–1932. Larmor published the complete Lorentz transformations if special relativity in the Philosophical Transactions of the Royal Society in 1897 some two years before Hendrik Lorentz (1899, 1904) and eight years before Albert Einstein (1905). Larmor predicted the phenomenon of time dilation, at least for orbiting electrons, and verified that the FitzGerald-Lorentz contraction (length contraction) should occur for bodies whose atoms were held together by electromagnetic forces. This however was all in the context of an aether theory of space-time.

(ii) Vector algebra method

Now we resolutely refuse to use choose axes at all.

We first dot equation (2.14) with $\mathbf{B}$ to obtain

$$m\ddot{\mathbf{r}}.\mathbf{B} = e\mathbf{E}.\mathbf{B}.$$

This can be integrated directly since $\ddot{\mathbf{r}}.\mathbf{B} = \dfrac{d^2(\mathbf{r}.\mathbf{B})}{dt^2}$:

$$\mathbf{r}.\mathbf{B} = (e/2m)\mathbf{E}.\mathbf{B}t^2 + at + b\,. \tag{2.19}$$

This is equivalent, in the coordinate-dependent method, to the z-equation (2.18).

What now? To be sure that no information is lost, we should really next cross equation (2.14) with $\mathbf{B}$. We would then have taken first the component of the equation parallel to $\mathbf{B}$ and subsequently the component perpendicular to $\mathbf{B}$. That would be a systematic approach. We could also dot with $\dot{\mathbf{r}}$, and integrate:

$$\tfrac{1}{2}\dot{\mathbf{r}}.\dot{\mathbf{r}} = e\mathbf{E}.\mathbf{r} + \text{constant}$$

giving an energy-like conservation equation, which may or may not be helpful (it isn't particularly helpful for present purposes).[25]

However, the easiest way forward in this particular case is to integrate the vector equation once directly, giving:

$$m\dot{\mathbf{r}} = e\mathbf{E}t + e\mathbf{r}\times\mathbf{B} + \mathbf{C}$$

where $\mathbf{C}$ is a vector constant of integration.[26] Now that we have an expression for $\dot{\mathbf{r}}$, we can substitute it into the right hand side of the equation of motion (2.14) to obtain an equation of the form (the details are getting messy, so the constant vectors are just called $\mathbf{A}_i$):

$$\begin{aligned} m\ddot{\mathbf{r}} =& \mathbf{A_1} + \mathbf{A_2}t + (e^2/m)(\mathbf{r}\times\mathbf{B})\times\mathbf{B} \\ =& \mathbf{A_1} + \mathbf{A_2}t + (e^2/m)\left((\mathbf{r}.\mathbf{B})\mathbf{B} - (\mathbf{B}.\mathbf{B})\mathbf{r}\right) \\ =& \mathbf{A_3} + \mathbf{A_4}t + \mathbf{A_5}t^2 - (e^2B^2/m)\mathbf{r} \\ =& -(e^2B^2/m)\mathbf{r} + \text{ other stuff} \end{aligned}$$

where we have used in the penultimate equation the expression (2.19) for $\mathbf{r}.\mathbf{B}$. This is just the vector simple harmonic motion equation (or three individual simple harmonic motion equations if we wrote it out in Cartesian coordinates) with additional forcing terms. The solution to this equation can more or less be written down:

$$\mathbf{r} = \mathbf{C_1}\cos\omega t + \mathbf{C_2}\sin\omega t + \text{Particular integral}$$

in agreement with what was obtained rather more easily in components.

End of example

2.4 Gravitational forces

2.4.1 Newton's universal law of gravitation

Newton's law of gravitation (published in *Principia* in 1687)[27] states that the gravitational force experienced by a particle of mass m_2 due to a particle of mass m_1 at distance r has magnitude

$$\frac{Gm_1m_2}{r^2}. \tag{2.20}$$

This is the inverse square law of gravitational attraction. The constant G in this expression is *Newton's gravitational constant*, aka the *universal gravitational constant* or just 'big G'. It has a

[25] It is not exactly conservation of energy because there is no potential for the Lorentz force in this case, and hence no potential energy. It is instead a statement about the work done by the force on the particle: the element of force due to the magnetic field is perpendicular to the velocity and hence does no work.

[26] Recall that $\dfrac{d(\mathbf{r}\times\mathbf{B})}{dt} = \dot{\mathbf{r}}\times\mathbf{B}$.

[27] There is some controversy about whom credit for this law should be attributed. Certainly, Hooke, Halley and Christopher Wren had all discussed it. What is not controversial is that Newton demonstrated that planets would move in ellipses, in agreement with observations, if moving under the influence of an inverse square law.

value of $6.6730010^{-11}\mathrm{m^3kg^{-1}s^{-2}}$. Note the dimensions: $\mathrm{L^3M^{-1}T^{-2}}$. Its value is quite hard to measure because gravitational forces are comparatively very weak.[28] For example, the ratio of the strength of the gravitational force between a proton and an electron to the strength of the electrostatic force between a proton and an electron (2.13) the same distance apart is

$$\frac{G \times m_{proton} m_{electron}}{q_{proton} q_{electron}/4\pi\epsilon_0} = \frac{(6.6 \times 10^{-11}) \times (1.6 \times 10^{-27}) \times (9 \times 10^{-31})}{1.6 \times 10^{-19}) \times (1.6 \times 10^{-19})/4\pi \times 8.8 \times 10^{-12}} \approx 10^{-39}$$

Combinations of the form GM_S and GM_E, where M_S and M_E are the mass of the sun and the Earth, respectively, are much easier to determine: they can be deduced from the period and radius of the orbits of the Earth and the Moon (see section 3.3).

The gravitational force between to particles is *central* (which means that it is directed from one particle towards the other) and attractive, so can be expressed in vector form as

$$\mathbf{F}_{12} = -\frac{Gm_1m_2}{r^3}\mathbf{r} \tag{2.21}$$

where $\mathbf{r}$ is the vector from particle 1 to particle 2 and $\mathbf{F}_{12}$ is the force exerted *by* the particle of mass m_1 *on* the particle of mass m_2. . In more general notation,

$$\mathbf{F}_{12} = -\frac{Gm_1m_2}{|\mathbf{r}_2 - \mathbf{r}_1|^3}(\mathbf{r}_2 - \mathbf{r}_1)\,.$$

As proved in section 2.1, all central forces are conservative. The potential for the force (2.21) is given by (recall that $\boldsymbol{\nabla} r^n = nr^{n-2}\mathbf{r}$)

$$\phi(r) = -\frac{Gm_1m_2}{r}; \qquad (\boldsymbol{\nabla}\phi = -\mathbf{F})$$

In more general notation,

$$\phi_{12}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{Gm_1m_2}{|\mathbf{r}_2 - \mathbf{r}_1|}.$$

We can take the gradient of this expression for the potential with respect to $\mathbf{r_2}$, regarding $\mathbf{r_1}$ as fixed, to obtain the force on particle 2 due to particle 1, or vice versa; the difference will only be minus sign (as expected from Newton's third law).

2.4.2 Important note

It is normal when considering gravitational potentials to omit the mass of the particle being acted on (the passive particle). Thus the gradient of the potential would give the acceleration of the passive particle not the force acting on it. I will distinguish between the two usages by using lower case ϕ for the potential which is equal to the potential energy of the particle (the gradient of the which gives the force), and upper case Φ for the potential more commonly used for gravitational and electric fields, the gradient of which gives the acceleration. Thus for a particle of mass M at the origin, the gravitational potential Φ is given by

$$\Phi(r) = -\frac{GM}{r}$$

whereas a particle of mass m moving in this potential would experience a force derived ('force = minus gradient of potential') from the potential function

$$\phi(r) = -\frac{GMm}{r}.$$

2.4.3 Addition of gravitational fields

Newtonian gravitational potentials are *linear* in the sense that the total potential due to two particles is just the sum of the potentials of the individual potentials. This is an observationally determined

[28] It was measured by Henry Cavendish in 1798 using a torsion balance.

result and does not hold for all types of potential.[29] Thus the potential at the point $\mathbf{r}$ due to point masses at points $\mathbf{r}_i$ $(i = 1,\ 2,\ \ldots)$ is

$$\Phi(\mathbf{r}) = -\sum_i \frac{Gm_i}{|\mathbf{r} - \mathbf{r}_i|}$$

and the total gravitational force on a particle of mass m at $\mathbf{r}$ is

$$-m\boldsymbol{\nabla}\Phi = -\sum_i \frac{Gm_i m(\mathbf{r} - \mathbf{r}_i)}{|\mathbf{r} - \mathbf{r}_i|^3}.$$

If all the masses are smeared out into a mass distribution with density $\rho(\mathbf{r})$, so that the mass in a volume dV of space is ρdV, the sums can be replace by a volume integrals[30] to obtain the total potential at the point $\mathbf{r}$:

$$\Phi(\mathbf{r}) = -\int \frac{G\rho(\mathbf{r}')\,dV'}{|\mathbf{r} - \mathbf{r}'|} \tag{2.22}$$

and the total gravitational force on a particle of mass m at $\mathbf{r}$ is

$$-\int \frac{Gm\rho(\mathbf{r}')\,(\mathbf{r} - \mathbf{r}')\,dV'}{|\mathbf{r} - \mathbf{r}'|^3}.$$

This is obtained by simply differentiating under the integral sign in (2.22), noting that $\mathbf{r}'$ is a dummy variable and is therefore a constant as far as differentiation with respect to $\mathbf{r}$ is concerned.

2.4.4 Gravitational field of a spherically symmetrical body

This is an important example: what we shall show is that the external gravitational field of a spherically symmetric body, such as a planet, of mass M is the same as that of a particle of mass M located at the centre of the body. **BUT** important though this result is, you should not regard the following calculation as being part of this course; it is really just an example of a volume integral as might be calculated in the Vector Calculus course, so you should stop reading this now, and come back to it when you are revising Vector Calculus. Though actually, the calculation is not very difficult.

We will demonstrate the result by evaluating the integral (2.22) to find the gravitational potential. Let the density of the body be $\rho(r)$ (which just depends on r, the distance from the centre, because the body is spherically symmetric), and let the radius of the body be a. We will calculate the gravitational field at a fixed point with position vector $\mathbf{R}$, a distance R from the centre, where $R \geq a$.

The first step is to choose coordinates. Obviously, we will use spherical coordinates, but the trick is to choose the polar direction $\theta = 0$ in the direction of $\mathbf{R}$. This means that for a position vector $\mathbf{r}$, the scalar product $\mathbf{R}.\mathbf{r} = Rr\cos\theta$. Further,

$$|\mathbf{R} - \mathbf{r}|^2 = (\mathbf{R} - \mathbf{r}).(\mathbf{R} - \mathbf{r}) = R^2 + r^2 - 2\mathbf{r}.\mathbf{R} = R^2 + r^2 - 2Rr\cos\theta.$$

[29] A more modern view (19th century — more modern than Newton) is that the forces of nature can be derived from potentials that satisfy Laplace's equation ($\nabla^2\phi = 0$). Since this equation is linear, solutions can be superposed. Einstein's equations for general relativity are non-linear and solutions cannot, in general, be added to obtain a new solution.

[30] You will come across volume integrals in the Vector Calculus course; we will not need to perform complicated integrals in this course.

Thus

$$
\begin{aligned}
\Phi(\mathbf{R}) &= -\int_{|\mathbf{r}|\le a} \frac{G\rho(r)dV}{\sqrt{R^2+r^2-2rR\cos\theta}} \\
&= -\int_0^a\int_0^\pi\int_0^{2\pi} \frac{G\rho(r)r^2\sin\theta d\phi d\theta dr}{\sqrt{R^2+r^2-2rR\cos\theta}} && \text{(remembering to put in the Jacobian)} \\
&= -\int_0^a\int_0^\pi \frac{2\pi G\rho(r)r^2\sin\theta d\theta dr}{\sqrt{R^2+r^2-2rR\cos\theta}} && \text{(doing the trivial } \phi\text{-integral)} \\
&= -\int_0^a\int_{-1}^1 \frac{2\pi G\rho(r)r^2 dr dc}{\sqrt{R^2+r^2-2rRc}} && \text{(setting } \cos\theta = c\text{)} \\
&= \int_0^a \frac{2\pi G\rho(r)r\left[\sqrt{R^2+r^2-2rRc}\right]_{c=-1}^{c=1} dr}{R} && \text{(doing the } c \text{ integral)} \\
&= -\int_0^a \frac{2\pi G\rho(r)r\left(|R+r|-|R-r|\right)dr}{R} && \text{(evaluating at } c=\pm1\text{)} \\
&= -\int_0^a \frac{4\pi G\rho(r)r^2 dr}{R} && \text{(using } r\le a\le R\text{)} \\
&= -\frac{GM}{R}
\end{aligned}
$$

as required.

End of example

2.5 Escape velocity

For a particle moving in a force field, the escape velocity is just the velocity that the particle must have to get out of the influence of the field; which normally means out to infinity. Often, one is thinking of projecting a particle from the surface of the Earth (say): the escape velocity tells you how fast you must project it for it to escape the Earth's gravitational pull.

For a general force field, the concept of escape velocity is not very helpful: the escape velocity would depend on the trajectory, and would not be possible to calculate without completely solving the equations of motion.

For a force field derived from a potential, such as a gravitational field, the concept is more useful because there is some chance that the escape velocity can be expressed in terms of the potential, without having to solve the equations of motion. If the particle has sufficient energy to overcome the potential it will escape. This is what is illustrated in the example on the motion of a particle in a cubic potential in section 2.1.

Even for a gravitational field, the concept only works well in the simplest case, the field of a single spherically symmetrical body such as (to good approximation) the Earth or the Sun. Even in the case of just two gravitating bodies, the escape velocity can depend critically on the direction of projection of the particle. For example, interplanetary probes use what is called the 'slingshot' effect to give the probe extra momentum by choosing the direction of projection so that the probe passes close to other planets. Voyager 1, which is now the furthest human-made object from Earth, is in the boundary zone between the Solar System and interstellar space. It gained the energy to escape the Sun's gravity completely by performing slingshot manoeuvres around Jupiter and Saturn. The energy gained was of course taken from the two planets, which perhaps slowed down[31] or moved further apart.

For a spherically symmetric planet of radius R, the gravitational potential at the surface is

$$-\frac{GM}{R}$$

[31] But not by much.

(see the example in section 2.5). Thus a particle of mass m projected with speed v from the surface has total energy E given by

$$E = \tfrac{1}{2}mv^2 - \frac{GMm}{R}$$

which is conserved. The potential energy of the particle if it escaped to infinity would be zero, so in order to have sufficient energy to escape,

$$E > 0,$$

i.e.

$$v > \sqrt{\frac{2GM}{R}} \equiv v_{\text{esc}}.$$

The minimum value of v, namely v_{esc}, is the escape velocity; or, more properly, the escape speed since it is independent of direction.

Clearly a particle that has less than this speed on projection cannot escape and will fall back to the point of projection. A particle that has at least this speed will escape and eventually (taking infinite time) reach infinity. This last statement is perhaps not quite obvious: clearly, the particle can reach infinity if it is projected radially outwards, because it could only turn round if its speed (and hence its kinetic energy) reduced to zero which, by conservation of energy, is impossible if $E > 0$. But what if it is projected tangentially? As we shall see in chapter 3, it then follows a parabolic or hyperbolic path, again out to infinity.

2.6 Three kinds of mass

This is an extended footnote: interesting, I hope, and relevant; but not strictly part of the course.

We can recognise three different sorts of mass that arise in Newtonian dynamics:

- *Inertial mass*, which occurs in Newton's second law:

$$\text{force} = \text{inertial mass} \;\times\; \text{acceleration}.$$

- *Passive gravitational mass*, which measures the response of a particle to a gravitational field. For example, at the surface of the Earth, the vertical force on a particle is given by

$$\text{passive gravitational mass} \times g.$$

- *Active gravitational mass*, which measures the magnitude of the gravitational field produced by a massive body.

All three kinds of mass occur simultaneously in the formula for the acceleration $\mathbf{a}$ of particle 1 of inertial mass $m_{\text{i}}^{(1)}$ and passive gravitational mass $m_{\text{p}}^{(1)}$ moving with acceleration $\mathbf{a}$ in the gravitational field of particle 2 of active gravitational mass $m_{\text{a}}^{(2)}$:

$$m_{\text{i}}^{(1)}\mathbf{a} = -\frac{Gm_{\text{p}}^{(1)}m_{\text{a}}^{(2)}\mathbf{r}}{r^3}$$

The fact that we only use one kind of mass, that is, we assume that the three apparently different kinds of mass are the same, needs explanation.

2.6.1 Equality of active and passive gravitational mass

According to the law of universal gravitation, the gravitational force on particle 1 due to particle 2, $\mathbf{F_{12}}$, is given by

$$\mathbf{F_{12}} = \frac{Gm_{\text{p}}^{(1)}m_{\text{a}}^{(2)}(\mathbf{r_2} - \mathbf{r_1})}{|\mathbf{r_2} - \mathbf{r_1}|^3}$$

and the gravitational force on particle 2 due to particle 1, $\mathbf{F_{21}}$ is given by

$$\mathbf{F_{21}} = \frac{Gm_{\text{a}}^{(1)}m_{\text{p}}^{(2)}(\mathbf{r_1} - \mathbf{r_2})}{|\mathbf{r_1} - \mathbf{r_2}|^3}\ .$$

Newton's third law demands that these forces are equal in magnitude, so we require

$$\frac{m_{\mathrm{a}}^{(1)}}{m_{\mathrm{p}}^{(1)}} = \frac{m_{\mathrm{a}}^{(2)}}{m_{\mathrm{p}}^{(2)}}$$

and furthermore that this relationship holds for all particles. Since the ratio of active and passive gravitational masses is equal for all particles, we can choose it to be unity (which would just involve scaling G).

2.6.2 Equality of inertial and gravitational mass

This is more difficult: we have so far in this course encountered no law or principle that would determine or even suggest a relationship between inertial and gravitational mass. Nevertheless, inertial and gravitational mass have been found in a number of celebrated experiments to coincide to a very high degree. It is, for example, what Galileo was trying to demonstrate by (supposedly) dropping objects from the top of the leaning tower of Pisa.

If you slide a particle down a slope and measure the acceleration a, you have

$$m_{\mathrm{i}} a = m_{\mathrm{p}} g \sin\theta,$$

where θ is the angle of the slope to the horizontal. If the acceleration is found to be the same for different particles then the only varying quantity in the above equation, namely the ratio $m_{\mathrm{p}}/m_{\mathrm{i}}$ must in fact be the same for the different particles and, as before, it can be normalised to one.

The Hungarian physicist Eötvös[32] spend much of his working life demonstrating the equality of inertial and gravitational mass. His method was to suspend two heavy spheres made from different material from a torsion balance, which consists of a horizontal rod suspended from a fixed point by a quartz fibre attached to its midpoint. The two spheres experience the Earth's gravitational force and also a centrifugal force due to the rotation of the Earth (see chapter 3), The spheres were arranged so that the rod was exactly horizontal, which means that the gravitational masses balanced exactly. If the inertial masses did not balance, the rod would rotate. It didn't.

This experiment was improved by Robert Dicke, an American physicist, using the effect of the Sun's gravitational field, which would have given a 24-hour periodic oscillation if the inertial mass and gravitational masses were inequivalent; this is extremely sensitive and established the equivalence to an accuracy of 1 part in 10^{12}.

The equivalence of inertial and gravitational mass (the *principle of equivalence*) is a fundamental pillar of modern physics; without it General Relativity, which interprets gravitational forces as fictitious forces (i.e. like centrifugal forces) due to motion in a curved space-time, would collapse.

[32] Vásárosnaményi Bárö Eötvös Loránd, 1848–1919; his surname is pronounced, roughly, utvush (u as in 'put').