Physics 303/573

Calculus in three dimensions

September 16, 2017

1 Mathematical preliminaries

I will begin by reviewing multivariate calculus, including the definitions of the gradient, curl, and divergence operators, and intuitive proofs of Stokes theorem and Gauss' theorem (divergence theorem).

1.1 Gradient

The gradient of a scalar function, that is, a function from three dimensions into the real numbers (actually, we can easily consider any number of dimensions) is defined by a Taylor expansion: if we know a f function at a point $\vec{r}$, then at a nearby point $\vec{r}+d\vec{r}$ we have

$$f(\vec{r}+d\vec{r}) := f(\vec{r}) + d\vec{r}\cdot\vec{\nabla}f(\vec{r}) + \mathcal{O}(dr^2) \tag{1.1}$$

where $\vec{\nabla}f$ is the *gradient* of f. It is a vector that points in the direction of greatest change of the function.

We can calculate the explicit form of the gradient in various coordinate systems. In Cartesian coordinates, $d\vec{r} = \hat{x}dx + \hat{y}dy + \hat{z}dz$ and

$$f(x+dx, y+dy, z+dz) = f(x,y,z) + dx\frac{\partial}{\partial x}f + dy\frac{\partial}{\partial y}f + dz\frac{\partial}{\partial z}f + ... \tag{1.2}$$

and hence (1.1) gives:

$$\vec{\nabla}f = \left(\hat{x}\frac{\partial}{\partial x} + \hat{y}\frac{\partial}{\partial y} + \hat{z}\frac{\partial}{\partial z}\right)f \tag{1.3}$$

whereas in cylindrical coordinates (recall section 2 in Lecture 2)

$$d\vec{r} = d(R\hat{R} + z\hat{z}) = \hat{R}dR + Rd\hat{R} + \hat{z}dz = \hat{R}dR + R\hat{\phi}d\phi + \hat{z}dz \tag{1.4}$$

and

$$f(R+dR,\phi+d\phi,z+dz)=f(x,y,z)+dR\frac{\partial}{\partial R}f+d\phi\frac{\partial}{\partial\phi}f+dz\frac{\partial}{\partial z}f+... \tag{1.5}$$

and hence (1.1) gives:

$$\vec{\nabla}f=\left(\hat{R}\frac{\partial}{\partial R}+\hat{\phi}\frac{1}{R}\frac{\partial}{\partial\phi}+\hat{z}\frac{\partial}{\partial z}\right)f \tag{1.6}$$

Finally, in spherical coordinates, (recall eq. (2.26) in Lecture 2)

$$d\vec{r}=d(r\hat{r})=\hat{r}dr+rd\hat{r}=\hat{r}dr+r(\hat{\theta}d\theta+\hat{\phi}\sin\theta d\phi) \tag{1.7}$$

and we find

$$\vec{\nabla}f=\left(\hat{r}\frac{\partial}{\partial r}+\hat{\theta}\frac{1}{r}\frac{\partial}{\partial\theta}+\hat{\phi}\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\right)f \tag{1.8}$$

If we integrate the gradient of a function along any path from one point to another, we find

$$\int_{\vec{a}}^{\vec{b}}d\vec{r}\cdot\vec{\nabla}f(\vec{r})=f(\vec{b})-f(\vec{a}) \tag{1.9}$$

This is just the fundamental theorem of calculus and follows from the definition of the integral as the limit of a sum and the definition (1.1) of the gradient; suppose we have a path from $\vec{a}$ to $\vec{b}$ made up of little segments $d\vec{r}_i$ such that

$$\vec{a}+\sum_{i=1}^{n}d\vec{r}_i=\vec{b} \tag{1.10}$$

Then

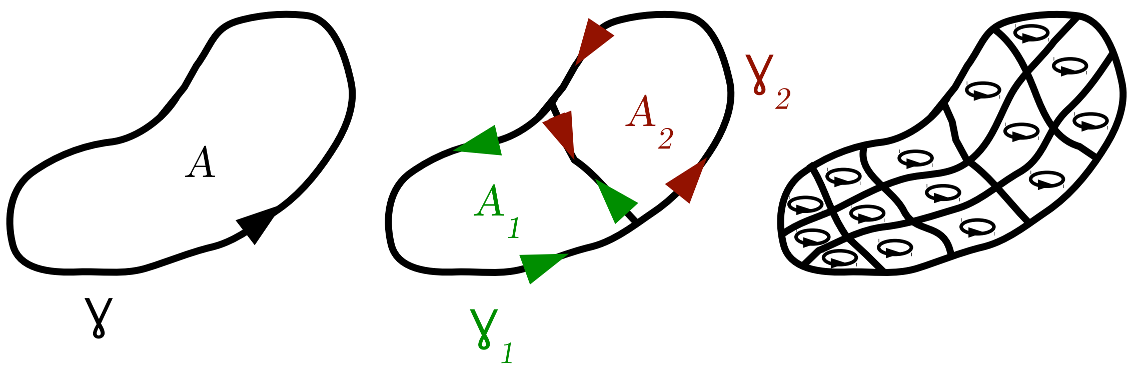
$$\begin{aligned}
\int_{\vec{a}}^{\vec{b}}d\vec{r}\cdot\vec{\nabla}f(\vec{r}) &= \lim_{n\to\infty}\sum_{i=1}^{n}d\vec{r}_i\cdot\vec{\nabla}f(\vec{a}+\sum_{j=1}^{i-1}d\vec{r}_j)\\
&= \lim_{n\to\infty}\left(\begin{array}{l} d\vec{r}_1\cdot\vec{\nabla}f(\vec{a})+d\vec{r}_2\cdot\vec{\nabla}f(\vec{a}+d\vec{r}_1)+d\vec{r}_3\cdot\vec{\nabla}f(\vec{a}+d\vec{r}_1+d\vec{r}_2)\\ +\ldots+d\vec{r}_{n-1}\cdot\vec{\nabla}f(\vec{b}-d\vec{r}_n-d\vec{r}_{n-1})+d\vec{r}_n\cdot\vec{\nabla}f(\vec{b}-d\vec{r}_n)\end{array}\right)\\
&= \lim_{n\to\infty}\left(\begin{array}{l} [f(\vec{a}+d\vec{r}_1)-f(\vec{a})]+[f(\vec{a}+d\vec{r}_1+d\vec{r}_2)-f(\vec{a}+d\vec{r}_1)]\\ +[f(\vec{a}+d\vec{r}_1+d\vec{r}_2+d\vec{r}_3)-f(\vec{a}+d\vec{r}_1+d\vec{r}_2)]+\ldots+\\ [f(\vec{b}-d\vec{r}_n)-f(\vec{b}-d\vec{r}_n-d\vec{r}_{n-1})]+[f(\vec{b})-f(\vec{b}-d\vec{r}_n)]\end{array}\right)\\
&= f(\vec{b})-f(\vec{a})
\end{aligned} \tag{1.11}$$

1.2 Curl

We have seen that the line integral of the gradient of a function along any path gives us just the difference of the function evaluated at the endpoints. In particular, it follows that the integral of the gradient along a closed path always gives 0. In a similar way, we can integrate an *arbitrary* vector field $\vec{v}(\vec{r})$ along an arbitrary path. In general, the value of the integral will depend on the path and will not vanish for a closed path. We define the *circulation* of the vector field $\vec{v}(\vec{r})$ through a region as the line integral along a closed path enclosing that region:

$$\mathrm{Circ}[\vec{v}(\vec{r}), A] = \oint_{\gamma=\partial A} \vec{v} \cdot d\vec{r} \tag{1.12}$$

where γ is a path around the boundary of the region A. In Figure 1a) such a region and the curve going around it are shown. The line integral of $\vec{v}$ along the boundary is the



a) b) c)

Figure 1: a) A region A bounded by an oriented curve γ. b) The region A is subdivided into two regions A_1 and A_2 bounded by oriented curves γ_1 and γ_2 respectively. c) The region A is further subdivided into many small regions.

circulation in A. In Figure 1b), A has been subdivided into two regions A_1 and A_2; notice that along their common boundary, the orientation of the bounding curves γ_1 and γ_2 is opposite, and thus the line integral cancels along this common boundary and we have

$$\oint_{\gamma=\partial A} \vec{v} \cdot d\vec{r} = \oint_{\gamma_1=\partial A_1} \vec{v} \cdot d\vec{r} + \oint_{\gamma_2=\partial A_2} \vec{v} \cdot d\vec{r} \tag{1.13}$$

whenever $A = A_1 \cup A_2$. As shown in Figure 1c), we can continue, and subdivide A into many little regions A_i with boundaries γ_i, and the sum of the line integrals around all the little curves γ_i is still the total integral over γ:

$$\oint_{\gamma=\partial A} \vec{v} \cdot d\vec{r} = \sum_{i=1}^{n} \oint_{\gamma_i=\partial A_i} \vec{v} \cdot d\vec{r} \tag{1.14}$$

We define the curl $\vec{\nabla} \times \vec{v}$ of the vector field $\vec{v}$ by

$$(\vec{\nabla} \times \vec{v}\,) \cdot \hat{n}_A = \lim_{A \to 0} \frac{1}{A} \oint_{\gamma=\partial A} \vec{v} \cdot d\vec{r} \tag{1.15}$$

where $\hat{n}_A$ is a unit vector perpendicular to the area A. This definition makes Stokes theorem obvious; we integrate the curl over the region A, and then use the definition (1.15) and the fact that the line integral over the boundary γ of A can be written as a sum of the line integrals over the boundaries of regions subdividing A (1.14) to find

$$\begin{aligned}
\int_A (\vec{\nabla} \times \vec{v}\,) \cdot \hat{n}_A dA &= \lim_{n \to \infty} \sum_{i=1}^{n} A_i \left(\frac{1}{A_i} \oint_{\gamma_i=\partial A_i} \vec{v} \cdot d\vec{r} \right) \\
&= \lim_{n \to \infty} \sum_{i=1}^{n} \oint_{\gamma_i=\partial A_i} \vec{v} \cdot d\vec{r} \\
&= \oint_{\gamma=\partial A} \vec{v} \cdot d\vec{r}
\end{aligned} \tag{1.16}$$

That is, the integral of the curl of a vector field over a region equals the line integral of the vector field around the boundary of the region:

$$\int_A (\vec{\nabla} \times \vec{v}\,) \cdot d\vec{A} = \oint_{\gamma=\partial A} \vec{v} \cdot d\vec{r} \tag{1.17}$$

where we use the notation $d\vec{A} = \hat{n}_A dA$ for the area element.

We now calculate the formula for the curl in Cartesian coordinates using the definition (1.15). We consider a little rectangular area element in the $x-y$ plane with area $A = dx\,dy$ as shown in Figure 2; $\hat{n}_A$ points along the z-axis, and hence we are calculating the z component of the curl. The line integral along the bottom of the curve gives us

$$\int_{(x,y)}^{(x+dx,y)} v_x(x', y) dx' = v_x(x, y)\, dx + \mathcal{O}(dx^2) \tag{1.18}$$

Then we integrate up along the right edge to get

$$\int_{(x+dx,y)}^{(x+dx,y+dy)} v_y(x + dx, y') dy' = v_y(x + dx, y)\, dy + \mathcal{O}(dy^2) \tag{1.19}$$

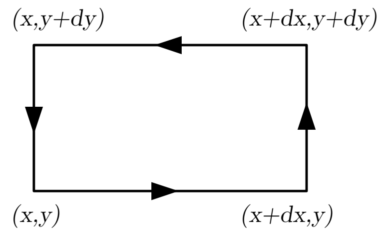


Figure 2: A small rectangle in the $x - y$ plane.

Integrating back along the top edge, we get

$$\int_{(x+dx,y+dy)}^{(x,y+dy)} v_x(x', y+dy)dx' = -v_x(x, y+dy)\, dx + \mathcal{O}(dx^2) \tag{1.20}$$

where the minus sign comes because we are integrating from $x + dx$ back to x. Finally, we close the contour to get

$$\int_{(x,y+dy)}^{(x,y)} v_y(x, y')dy' = -v_y(x, y)\, dy + \mathcal{O}(dy^2) \tag{1.21}$$

Adding this all up, we find

$$\begin{aligned} \oint_{\gamma=\partial A} \vec{v} \cdot d\vec{r} &= (v_y(x+dx, y) - v_y(x, y))dy - (v_x(x, y+dy) - v_x(x, y))dx + \mathcal{O}(dx^2, dy^2) \\ &= \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right) dx\, dy + \mathcal{O}(dx^2, dy^2) \qquad (1.22) \end{aligned}$$

and hence, using $A = dx\, dy$ and the definition (1.15), we find

$$(\vec{\nabla} \times \vec{v})_z = \frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y} \tag{1.23}$$

We can repeat this calculation for other orientations of the rectangle, and find the remaining components; the final answer is:

$$\vec{\nabla} \times \vec{v} = \hat{z}\left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right) + \hat{x}\left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right) + \hat{y}\left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right) \tag{1.24}$$

The same method can be used to compute the curl in other coordinate systems.

Since the line integral of the gradient of a function around a closed loop always vanishes, Stokes theorem implies that the curl of a gradient vanishes; it is very easy to check this in Cartesian coordinates, *e.g.*,

$$(\vec{\nabla} \times \vec{\nabla} f)_z = \frac{\partial}{\partial x}\frac{\partial f}{\partial y} - \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \partial x} - \frac{\partial^2 f}{\partial y \partial x} = 0 \tag{1.25}$$

There is one subtlety: it is possible to have a "vector"-field that isn't quite defined everywhere such that the curl vanishes everywhere except at the points where vector-field is not well defined, and then the line integral (circulation) will vanish if it doesn't enclose the bad points, but will not vanish is the path circles the bad points. A simple example is given by the "vector"

$$\vec{v} = \frac{\hat{y}\, x - \hat{x}\, y}{x^2 + y^2} \tag{1.26}$$

which is curl-free everywhere except the origin, and has no circulation for a path doesn't enclose the origin and has circulation 2π for a path that circles the origin counterclockwise once. The easiest way to prove this is to show that

$$\vec{v} = \vec{\nabla}\theta = \frac{\hat{\theta}}{R} \tag{1.27}$$

Here R, θ are cylindrical coordinates. Since we have written the vector $\vec{v}$ as a gradient, its curl vanishes everywhere except at $R = 0$, where it is undefined.

1.3 Divergence

Another property of a vector field, such as the velocity as a function of the position in a fluid, is the flux: the total amount flowing through a surface. We write

$$\Phi = \int_A \vec{v} \cdot \hat{n}_A dA = \int_A \vec{v} \cdot d\vec{A} \tag{1.28}$$

where as in (1.15), $\hat{n}_A$ is a unit vector perpendicular to the area A. The flux through a closed surface is the net amount of the vector field flowing out of or in to the enclosed volume. As shown in Figure 3, if we subdivide the volume V with an internal surface A_{12} into two volumes V_1 and V_2, the flux flowing out of V_1 flows into V_2, and hence the sum of the fluxes is the total flux:

$$\Phi(V) = \Phi(V_1) + \Phi(V_2) \tag{1.29}$$

We can keep subdividing the volume into smaller and smaller pieces, and define the divergence $\vec{\nabla} \cdot \vec{v}$ by the limit

$$\vec{\nabla} \cdot \vec{v} = \lim_{V \to 0} \frac{1}{V} \int_{A=\partial V} \vec{v} \cdot d\vec{A} \tag{1.30}$$

where A is the surface bounding the volume V. Gauss' theorem (also called the divergence theorem) is now obvious: The total flux flowing out through the surface of a volume is the integral of the divergence over the volume. To prove this, we write the total flux as the

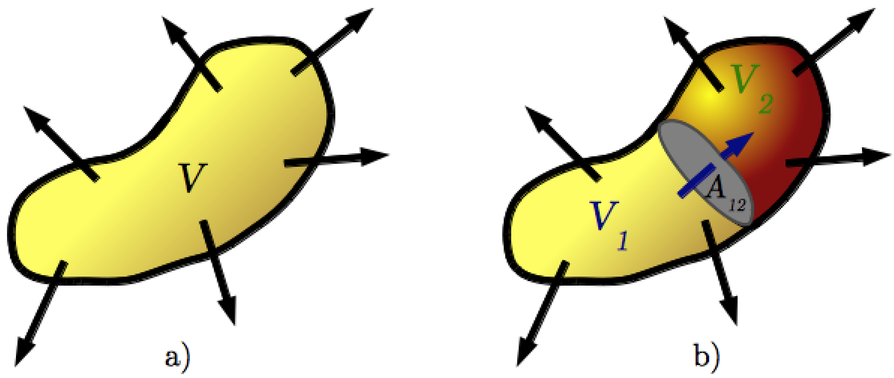


Figure 3: a) The flux from flowing through a surface around a volume V. b) The flux flowing through the surfaces of two volumes V_1 and V_2; on the interface A_{12} between the two, whatever flows out of V_1 flows into V_2.

sum of the fluxes of an ever finer subdivision of the volume V into smaller and smaller volumes V_i and use the definition (1.30):

$$\begin{aligned}
\Phi(V) \;=\; \int_{A=\partial V} \vec{v}\cdot d\vec{A} \;&=\; \lim_{i\to\infty}\sum_i \Phi(V_i) \\
&=\; \lim_{i\to\infty}\sum_i \int_{A_i=\partial V_i} \vec{v}\cdot d\vec{A} \\
&=\; \lim_{i\to\infty}\sum_i V_i\left(\frac{1}{V_i}\int_{A_i=\partial V_i} \vec{v}\cdot d\vec{A}\right) \\
&=\; \lim_{i\to\infty}\sum_i V_i\left(\vec{\nabla}\cdot\vec{v}\right) \;=\; \int_V \vec{\nabla}\cdot\vec{v}\,dV
\end{aligned} \tag{1.31}$$

and hence we have Gauss' theorem:

$$\Phi(V) = \int_{A=\partial V} \vec{v}\cdot d\vec{A} = \int_V \vec{\nabla}\cdot\vec{v}\;dV \tag{1.32}$$

Again, it is easy to compute the divergence in specific coordinate systems. Imagine a infinitesimal cube as shown in Figure 4. We can assume that the vector field $\vec{v}(\vec{r})$ is approximately constant on any given face of the cube.

The flux through the top is $v_z(\bar{x},\bar{y},z+dz)dxdy$ where as the flux through the bottom is $-v_z(\bar{x},\bar{y},z)dxdy$ where $\bar{x},\bar{y}$ are the average value of x,y on the top and bottom faces. Similarly, we have the flux through the left face is $v_x(x+dx,\bar{y},\bar{z})dydz$ and through the right face is $-v_x(x,\bar{y},\bar{z})dydz$; finally the flux through the back face is $v_y(\bar{x},y+dy,\bar{z})dxdz$

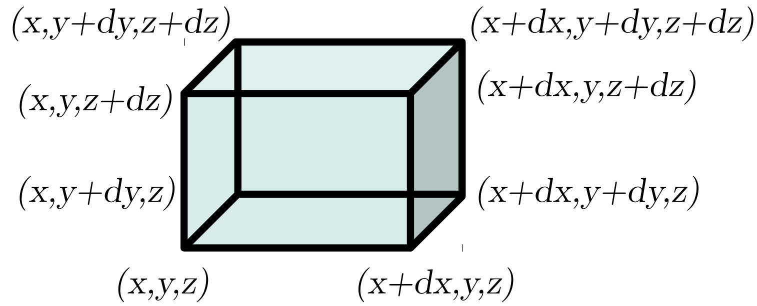


Figure 4: A little Cartesian cube for calculating the divergence.

whereas the flux through the front face is $-v_y(\bar{x}, y, \bar{z})dxdz$. Adding this all up and keeping only the first nontrivial terms in the Taylor expansion, we find:

$$\begin{aligned}
&v_z(\bar{x}, \bar{y}, z+dz)dx\,dy - v_z(\bar{x}, \bar{y}, z)dx\,dy + v_x(x+dx, \bar{y}, \bar{z})dy\,dz - v_x(x, \bar{y}, \bar{z})dy\,dz \\
&\qquad + v_y(\bar{x}, y+dy, \bar{z})dx\,dz - v_y(\bar{x}, y, \bar{z})dx\,dz \\
&= [v_z(\bar{x}, \bar{y}, z+dz) - v_z(\bar{x}, \bar{y}, z)]dx\,dy + [v_x(x+dx, \bar{y}, \bar{z}) - v_x(x, \bar{y}, \bar{z})]dy\,dz \\
&\qquad + [v_y(\bar{x}, y+dy, \bar{z}) - v_y(\bar{x}, y, \bar{z})]dx\,dz \\
&= \left(\frac{\partial v_z}{\partial z} + \frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y}\right) dx\,dy\,dz \qquad (1.33)
\end{aligned}$$

As the volume of the little cube is $dx\,dy\,dz$, the definition (1.30) implies that in Cartesian coordinates,

$$\vec{\nabla} \cdot \vec{v} = \left(\frac{\partial v_x}{\partial x} + \frac{\partial v_y}{\partial y} + \frac{\partial v_z}{\partial z}\right) \qquad (1.34)$$

2 Work

We now turn to a physical application of the mathematical concepts developed above.

In more than one dimension, we write Newton's laws using vectors:

$$\vec{F} = m\vec{a} = m\dot{\vec{v}} = m\ddot{\vec{r}} \qquad (2.35)$$

The time rate of change of the kinetic energy of an object moving in more than one dimension is

$$\dot{T} = \frac{d}{dt}\left(\frac{1}{2}m\vec{v} \cdot \vec{v}\right) = m\dot{\vec{v}} \cdot \vec{v} = m\vec{a} \cdot \vec{v} = \vec{F} \cdot \vec{v} \qquad (2.36)$$

where we use Newton's laws of motion in the final step. We can integrate this along a path γ to find:

$$T_2 - T_1 = \int_\gamma \vec{F} \cdot \vec{v} dt = \int_\gamma \vec{F} \cdot \vec{dr} \tag{2.37}$$

This change in the kinetic energy is called the work done by the force. Notice that we are integrating the component of the force along the path; in particular, if the force is perpendicular to the path, then the force does no work at all.

If we specify a path parametrically by giving $\vec{r}(s)$ for $s = s_i ... s_f$, then the line integral reduces to an ordinary integral:

$$\int_\gamma \vec{F} \cdot \vec{dr} = \int_{s_i}^{s_f} \vec{F} \cdot \left(\frac{d}{ds} \vec{r}(s) \right) ds \tag{2.38}$$

In (2.37), we have the special case where the parameter s is just the time: $s = t$.

In general, the work will depend on the path we follow, and in particular, can be nonvanishing even for a closed path that returns to where it started. In this case, we say the force is *not* conservative. Conversely, if the work done depends only on the initial and final position, or equivalently vanishes for any closed path, we say the force *is* conservative. This is much less trivial than for one-dimensional motion, where any force $F(x)$ that depends only on the position is automatically conservative. In terms of the mathematical notions introduced above, a conservative force has zero circulation (1.12) and in particular, zero curl (1.24).

It is easy to write down forces that are not conservative; for example

$$\vec{F} = \frac{1}{2}(x\hat{y} - y\hat{x}) = \frac{1}{2} r\hat{\theta} \tag{2.39}$$

spins around the z-axis and has curl $\hat{z}$. The work done by this force on a particle that moves along a circle of radius r is just the circulation of $\vec{F}$ and equals πr^2 (see Figure 5a), the work done along the contour shown in Figure 5b) is $\frac{\pi}{2}(b^2 - a^2)$, and the work done along the contour shown in Figure 5c) is ab. This is most easily computed using Stokes theorem, since the curl is perpendicular to the area enclosed by the contours and has magnitude 1, and so the work is just the area enclosed by the contour, but it is not hard to do the line integrals directly. In 5b), the segments of the contours along the x-axis do not contribute, as the component of the force along those segments vanishes. Similarly, in 5c), the segments of the contours along the x-axis and the along the y-axis do not contribute.

A conservative force can *always* be written as the gradient of a potential function $V(\vec{r})$: if we define

$$\vec{F} = -\vec{\nabla} V(\vec{r}) \tag{2.40}$$

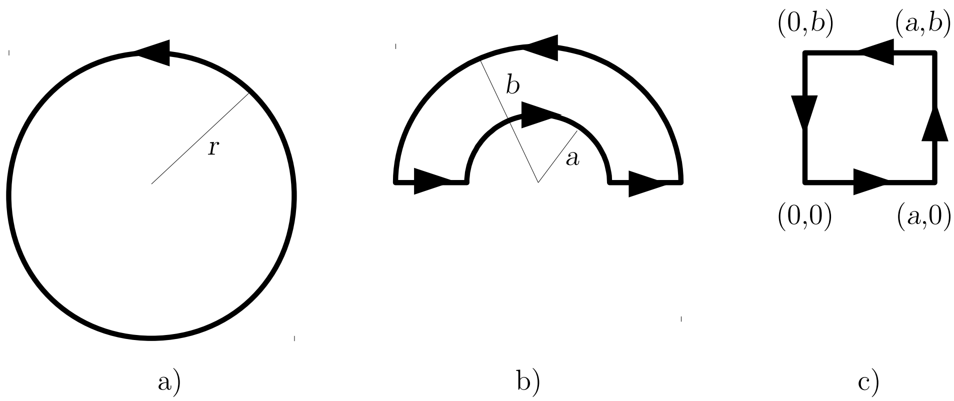


Figure 5: Three paths for calculating the work done by the force (2.39).

then

$$T(\vec{r}_2) - T(\vec{r}_1) = \int_{r_1}^{r_2} \vec{F} \cdot \vec{dr} = V(\vec{r}_1) - V(\vec{r}_2) \tag{2.41}$$

and hence the total energy $E = T + V$ doesn't change:

$$E(\vec{r}_2) = T(\vec{r}_2) + V(\vec{r}_2) = T(\vec{r}_1) + V(\vec{r}_1) = E(\vec{r}_1) \tag{2.42}$$

we say that any quantity that is constant in time is conserved: