Fluxes, Surface Integrals and Conservation Laws

February 1, 2016

Overview

These notes cover the following:

- Rate of transport through a surface and flux
- Vector fields and how to visualize them
- How to evaluate surface integrals
- Conservation of mass for a continuum — integral formulation

Rate of mass transfer: a simple case

We have seen how to describe the spatial distribution of mass in a continuum through a density field $\rho(x, y, z, t)$, and how to compute the mass contained in a given volume V from the density field by integrating,

$$M = \int_V \rho \, \mathrm{d}V.$$

Mass is a conserved quantity, and changes in M can therefore not happen without accounting for how mass got into our out of the volume V. This means we have to figure out how to compute the rate at which mass crosses the boundary of the volume.

To develop an intuition, we can look at a simple case. Assume you have a straight pipe or channel, for instance an idealized stream or river. Let the cross-sectional area of the channel be S, and assume that a fluid flows down the channel at a spatially uniform velocity, constant v pointing in the down-channel direction, and that the density ρ of the channel is also spatially uniform and constant in time. At what rate does mass pass through the cross-section S?

The answer clearly needs to have dimensions of mass over time, but you should not be tempted to answer this question by dimensional analysis alone — because there is no guarantee this will give you the right answer.[1] The simplest way to compute this is to look at the mass M that passes through S in some time interval T, and define the rate as M/T.

To compute the mass M that passes through, it makes sense to compute the volume that passes through S first. As the fluid travels a distance vT in time T, the body of water that passes through S is a prism-like volume[2] with base area S and an edge length vT. The base area S is at right angles to that edge, and the volume that passes through is therefore

$$SvT$$

The mass contained in the volume is density times volume, or $M = \rho SvT$. If we denote the rate at which mass passes through S by Q, we have

$$Q = \frac{M}{T} = \frac{\rho SvT}{T} = \rho vS.$$

Note that we can also define a rate at which mass passes through per unit area of channel-cross-section. This simply amounts to dividing by S, giving $Q/S = \rho v$. This measure of transport will later be called a *flux*. The advantage of this measure is that it does not depend on the cross-section that we are looking at.

In what hopefully begins to look like a pattern, we have another example of proportionality above: velocity and density being equal, the rate of mass transfer is proportional to the cross-sectional area of the channel we are looking at. The flux ρv is simply the constant of proportionality.

Transfer through oblique surfaces

We deliberately made the cross-sectional area S perpendicular to the channel above. In general, we may be interested in mass passing obliquely through a surface. Consider placing a different cross-section S' into the channel, this one still planar but at an angle to the flow direction.[3] Can we compute the rate at which mass passes through S' in a similar way?

Again, we should look at the volume that passes throuhg S' in some time T. Once more, the fluid travels a distance vT. The volume passes through is again a prism-like shape, now with base area S' and again with edge length vT. The edge is however no

[1] Computations done by dimensional analysis are only ever unique up to a numerical factor in any case.

[2] Strictly speaking, a 'prism' would require S to be a polygon

[3] There is nothing special about the notation S', and there are no hidden meanings like derivatives lurking here. We are simply using a prime to make a symbol that looks similar to the original cross-sectional area symbol S, but which is not completely the same.

longer perpendicular to the base. When computing the volume of such a prism-like shape, we again have a formula of

$$\text{volume} = \text{base} \times \text{height}.$$

The height is however no longer the edge length, but the height measured at right angles to the base. In this case, the height can be seen as the projection of the edge length vT onto the direction perpendicular to S'. This direction is also called the *normal* direction, and a vector that points in that direction is a *normal* to S'. If the angle between the velocity vector and the normal direction is θ, then the height of the prism-like surface is

$$vT\cos(\theta)$$

and the volume is $S'vT\cos(\theta)$. As before, mass is density times volume, so $M = \rho v S'\cos(\theta)T$, and the rate of mass transfer is M/T, or

$$Q = \rho v \cos(\theta) S'$$

Actually trying to figure out the angle θ between the flow direction and the normal direction is not a particular useful exercise. Instead, we can write the combination $v\cos(\theta)$ in the language of vectors if we define the idea of a *unit normal* to S': We let $\hat{\mathbf{n}}$ be a unit vector perpendicular to S', which is sraightforward to do as we have assumed that S' is planar. If we also define a vector velocity $\mathbf{v}$ (so that v above is the magnitude of $\mathbf{v}$), then we have

$$\mathbf{v} \cdot \hat{\mathbf{n}} = |\mathbf{v}||\hat{\mathbf{n}}|\cos(\theta)$$

But $\hat{\mathbf{n}}$ is a unit vector, so $|\hat{\mathbf{n}}| = 1$, and $|\mathbf{v}| = v$. Therefore

$$\mathbf{v} \cdot \hat{\mathbf{n}} = v\cos(\theta)$$

and

$$Q = \rho \mathbf{v} \cdot \hat{\mathbf{n}} S'. \tag{1}$$

This suggests that we should generalize the flux that we considered above to be a vector $\rho\mathbf{v}$. Again, the rate of mass transfer Q is proportional to S'. Because Q is a scalar quantity (mass transferred from one side of the surface S' to the other per unit time), the constant of proportionality cannot be the vector flux $\rho\mathbf{v}$. Instead, it is the scalar product of flux and unit normal $\hat{\mathbf{n}}$ that gives us the constant of proportionality.

The fact that the direction of the surface relative to the velocity vector features in the calculation should make a lot of sense. Imagine a surface S' that lies parallel to the flow direction $\mathbf{v}$. Mass would then flow tangentially to S' but never actually cross it. This is indeed the case in our calculation: If $\mathbf{v}$ is tangential to the surface S', then it is perpendicular to the normal $\hat{\mathbf{n}}$, and $\rho\mathbf{v} \cdot \hat{\mathbf{n}} = 0$.

Note 1 *There is an interesting note about the relationship between S' and S here. Presumably, as the flow of the stream is steady and uniform, no mass is being piled up between the surfaces S' and S, so the rates of mass transfer through them must be the same. We computed these as $\rho v S$ and $\rho v \cos(\theta) S'$. If these are equal, we must have*

$$S = S' \cos(\theta).$$

There are actually *two* possible choices of normal direction, corresponding to two unit normal vectors that are equal and opposite. Implicitly, we have chosen the one that makes an acute angle with the velocity vector $\mathbf{v}$ above, because we wanted to calculate a positive rate of mass transfer. Had we chosen the normal vector making an obtuse angle, we would have compute Q with the same magnitude, but negative in sign.

There is a straightforward interpretation of this, which will become important later: Mass transfer occurs *from* one side of S' *to* another. Therefore, although Q is not a vector, it does have a sense of direction. The choice of $\hat{\mathbf{n}}$ should point towards the side of S' toward which we want to compute the rate of mass transfer. If that rate of mass transfer turns out to be negative, we know that actual mass transfer is happening at the same rate but in the opposite direction. We can take this to be the definition of a negative rate of mass transfer.

Unsteady, non-uniform flow through curved surfaces: surface integrals

The calculations above were relatively simple — involving only some basic geometry and vector manipulation — because the cross-section S' was taken to be planar, and ρ and $\mathbf{v}$ were uniform in space and constant in time. These are assumptions that are unlikely to apply under most circumstances. In general, both velocity and density may vary in space and time, and we may need to compute rates of mass transfer across curved surfaces.

The way to deal with this should not come as a surprise. Assume we know the surface S as well as the velocity field $\mathbf{v}(x, y, z, t)$ and density field $\rho(x, y, z, t)$. We can then split the surface S into many small parts δS (called *surface elements*) and look at mass transfer δm through each in a small period of time from t to $t + \delta t$. Doing so allows us to treat ρ and $\mathbf{v}$ as approximately constant on each surface element in the time interval δt. Moreover, small surface elements can be treated as approximately planar, which allows us to identify a normal $\hat{\mathbf{n}}$ that can be treated as constant on each surface element.

Following the discussion at the end of the last section, we have to decide a definite direction for that normal vector $\hat{\mathbf{n}}$, given by the side of the surface S *towards* which the computed mass transfer occurs. For every surface element δS, the normal $\hat{\mathbf{n}}$ needs to point towards that side, even if $\mathbf{v}$ points in the opposite direction. Of course, it is then possible that computed rate of mass transfer may be negative. This simply means that, in practice, there is net mass transfer of the same magnitdue but towards the opposite side of S.

The calculation for how much mass δm passes through each δS in δt works in exactly the same way as for the oblique surface S' above, except that many δ's will appear.

We again look at the volume that passes through δS in the direction of the normal vector $\hat{\mathbf{n}}$ in time δt. That volume is still approximately a prism-like shape with size given by base area times height (figure 1). The base area is now δS, and height again cannot be identified as the side length $v\delta t$ because velocity $\mathbf{v}$ may not be perpendicular to the surface δS. Instead, the height of the 'prism' is $v\delta t\cos\theta$, where θ is the angle indicated in figure 1. So we have a volume whose size is approximately

$$\delta S v\delta t\cos\theta = \delta S\mathbf{v}\cdot\hat{\mathbf{n}}\delta t$$

and mass

$$\delta m \approx \rho v\cos\theta\delta S\delta t. \tag{2}$$

If $\hat{\mathbf{n}}$ and $\mathbf{v}$ form an obtuse angle and their dot product is negative, then the volume we have computed is negative. A 'negative volume' and hence a negative amount of mass has passed through δS in the direction of $\hat{\mathbf{n}}$, which means a positive volume and mass has passed in the opposite direction.

The amount of mass δM that passes through the whole surface S in δt is then found by adding the masses that pass through all the δS's:

$$\delta M = \sum_{\delta S}\delta m \approx \sum_{\delta S}\rho\mathbf{v}\cdot\hat{\mathbf{n}}\delta S\delta t.$$

The rate of mass transfer is then found by dividing by δt:

$$Q = \lim_{\delta t\to 0}\frac{\delta M}{\delta t} \approx \sum_{\delta S}\rho\mathbf{v}\cdot\hat{\mathbf{n}}\delta S.$$

The total rate of mass transfer is given by summing the normal components of flux over all the surface elements δs. Of course, implicit in our discussion is that the δS must be small, and therefore that we are really interested in the limit $\delta S\to 0$, in which case the approximations above become exact. The sum then becomes a *surface integral*, which we denote by

$$Q = \int_S \rho\mathbf{v}\cdot\hat{\mathbf{n}}\,\mathrm{d}S.$$

The subscript 'S' on $\int_S$ indicates the surface over which we are summing.

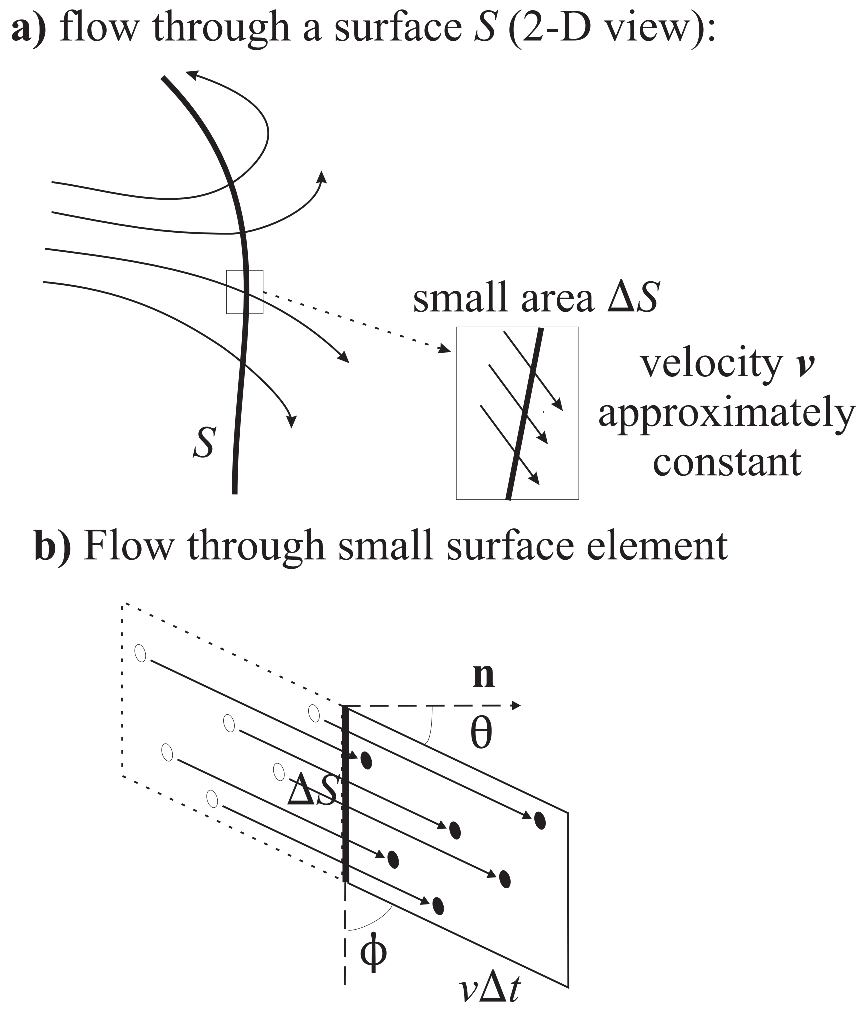

Figure 1: Calculating the mass of fluid that passes through S in time δt by splitting S into small surface elements δS (panel a). On each δS, velocity and density are approximately constant, and δS is approximately flat. The volume of fluid that passes through δS in δt is given by the normal component of velocity, $\mathbf{v} \cdot \hat{\mathbf{n}}$ (panel b).

Flux

We already introduced $\rho\mathbf{v}$ as a 'flux' that gives the constant of proportionality between rate of mass transfer and surface area in (1). As we have seen above, flux is in general a *field*, as both ρ and $\mathbf{v}$ can depend on position and time. Flux is often given its own symbol $\mathbf{q}$, and there can be fluxes associated with conserved quantities other than mass, so $\mathbf{q} = \rho\mathbf{v}$ should be called 'mass flux'.

When $\mathbf{q}(x, y, z, t)$ depends on position and time, mass transfer is no longer proportional to time and surface area for large time intervals or large surface areas. From (2) we see that the proportionality however still works for small surface areas and short times elapsed. The amount of mass δm that passes through a small area δS in a small amount of time δt does satisfy

$$\delta m \propto \delta S \delta t$$

and the constant of proportionality is $\mathbf{q} \cdot \hat{\mathbf{n}}$, where the normal vector $\hat{\mathbf{n}}$ is determined by the orientation of δS. Note that $\mathbf{q} \cdot \hat{\mathbf{n}}$ is the *normal component* of flux; any part of $\mathbf{q}$ that is tangential to δS does not contribute to mass transfer through the surface. We have

$$\delta m = \mathbf{q} \cdot \hat{\mathbf{n}} \delta S \delta t.$$

Remember that we had a similar argument for mass content in a small volume δV, where we had $\delta m = \rho \delta V$. This allows density at a point (x, y, z) to be defined in terms of the mass δm in a small volume δV centered on that point through

$$\rho(x, y, z, t) \approx \frac{\delta m}{\delta V},$$

the limit of small δV turning the approximation into an equality. The advantage of this definition is that it uses measurable quantities like mass δm and volume δV.

Suppose we would like to come up with a similar definition for flux $\mathbf{q}$ in terms of the mass δm transferred through a small surface δS centered on (x, y, z) in time δt, these again being potentially measurable quantities. This would allows to generalize eventually from mass flux, where we have the formula $\mathbf{q} = \rho\mathbf{v}$ to the fluxes of other physical quantities. From above, we have

$$\mathbf{q} \cdot \hat{\mathbf{n}} = \frac{\delta m}{\delta S \delta t},$$

which could almost serve as a definition of $\mathbf{q}$, saying that flux is the mass transferred over time elapsed and size of area, in the limit of a short time interval elapsed and a small surface area element. The problem is the dot product with the unit vector $\hat{\mathbf{n}}$.

Imagine that we are located at a point (x, y, z) and we can measure the mass δm that passes through a small area element δS in δt, but that we can rotate δS into whatever orientation we like. This means we can alter $\hat{\mathbf{n}}$ at will. In general, we have

$$\frac{\delta m}{\delta S \delta t} = \mathbf{q} \cdot \hat{\mathbf{n}} = |\mathbf{q}||\hat{\mathbf{n}}| \cos(\theta) = |\mathbf{q}| \cos(\theta) \leq |\mathbf{q}|$$

so the magnitude of the actual flux is greater than or equal to $\delta m/(\delta S \delta t)$,

$$|\mathbf{q}| \geq \frac{\delta m}{\delta S \delta t}.$$

However, when $\hat{\mathbf{n}}$ is parallel to the direction of flow $\mathbf{q}$, then $\cos(\theta) = 1$ and we get the equality

$$|\mathbf{q}| = \frac{\delta m}{\delta S \delta t}.$$

The magnitude of flux is therefore the maximum that the ratio $\delta m/(\delta S \delta t)$ can reach when the orientation of the surface element is changed. In other words,

$$|\mathbf{q}| = \frac{\text{maximum } \delta m \text{ transferred through } \delta S \text{ in } \delta t \text{ as } \delta S \text{ is rotated through all orientations}}{\delta S \delta t},$$

with the direction of $\mathbf{q}$ given by the normal to δS when that maximum is attained.

Note 2 *There are other naming conventions in some areas of physics; the one we have adopted here is common in continuum physics (more usually called 'continuum mechanics'). In electromagnetism, 'flux' is used to refer to the integral*

$$Q = \int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S,$$

and which case $\mathbf{q}$ *is referred to as the* flux density. *In our case, the variable* Q *will be referred to as the rate of transfer, and* $\mathbf{q}$ *as the flux.*

Specifying and visualizing scalar and vector fields

We have now seen two types of field, namely scalar fields like $\rho(x, y, z, t)$, and vector fields like $\mathbf{q}(x, y, z, t)$ or $\mathbf{v}(x, y, z, t)$. It is easy to see how to specify a scalar field; it is simply an ordinary function of position (x, y, z) and time t. What about a vector field?

A vector is most easily specified through its components, as $\mathbf{a} = a_x\mathbf{i} + a_y\mathbf{j} + a_z\mathbf{k}$ or $\mathbf{b} = \mathbf{i} + 2\mathbf{j} - \mathbf{k}$. For a vector field, the important aspect is that *each* of its components is a function of position (x, y, z)

$$\mathbf{v}(x, y, z, t) = v_x(x, y, z, t)\mathbf{i} + v_y(z, y, x, t)\mathbf{j} + v_z(x, y, z, t)\mathbf{k}. \tag{3}$$

Example 1 *For instance, if*

$$\mathbf{v} = (y - x)\mathbf{i} - (x + y - z)\mathbf{j} + \mathbf{k},$$

we have $v_x(x, y, z) = y - x$, $v_y(x, y, z) = -x - y + z$, $v_z(x, y, z) = 1$

There is no contradiction in the x-component v_x depending not only on x but also on y and z. Remember that the x-component of the vector $\mathbf{v}$ is the part of $\mathbf{v}$ that points in the direction of the x-axis. The coordinates (x, y, z) simply tell you the location at which you compute that vector. They are the components of the *position vector*

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$$

but not of the velocity vector $\mathbf{v}$ at that position.

Note 3 *The position vector of a point (x, y, z) is the vector that connects the origin to that point. We will always denote the position vector by $\mathbf{r}$. The length of the position vector is*

$$|\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}.$$

Physically, this is the distance from the origin to the point (x, y, z), and we usually denote this by $r = |\mathbf{r}| = \sqrt{x^2 + y^2 + z^2}$.

Sometimes we also need a unit vector in the direction from the origin to (x, y, z); you can think of this as a 'radial' unit vector as it points radially outwards from the origin. We would denote such a unit vector by

$$\hat{\mathbf{r}} = \frac{\mathbf{r}}{r} = \frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{\sqrt{x^2 + y^2 + z^2}}.$$

The dependence of $\mathbf{v}$ on (x, y, z) is easiest understood if we try to visualize a vector field. The way this is normally done is by defining a grid of points (x, y, z), usually evenly spaced and including the origin. Each of these points corresponds to a different combination of x, y and z. At each point, we can compute the components $v_x(x, y, z)$, $v_y(x, y, z)$ and $v_z(x, y, x)$, and therefore compute the vector $\mathbf{v}$ at that point. That vector is then drawn as an arrow, starting at that point, so the tail of the arrow indicates the location it refers to. It is possible that the vectors computed this way have large components that would make the arrows cross each other. This is usually avoided by scaling each arrow with a common scaling factor (for instance by reducing the length of each arrow by a factor of 10 before plotting them).

Exercise 1 *Sketch the following vector fields in this way:*

1. $\mathbf{v} = (y - x)\mathbf{i} - (x + y)\mathbf{j}$
2. $\mathbf{v} = -y\mathbf{i} + x\mathbf{j}$
3. $\mathbf{v} = -x\mathbf{i} + y\mathbf{j}$
4. $\mathbf{v} = x\mathbf{i} + xy\mathbf{j}$

A second way of visualizing vector fields is through *streamlines*. This is particularly easy to understand if we treat $\mathbf{v} = v_x\mathbf{i} + v_y\mathbf{j} + v_z\mathbf{k}$ as an actual velocity field in, say, a fluid. To make life easier, assume we are in two dimensions, so

$$\mathbf{v} = v_x(x, y)\mathbf{i} + v_y(x, y)\mathbf{j}$$

Assume that the velocity field is 'steady', meaning v_x and v_y do not depend on time.

Now imagine that we want to follow a particle immersed in that material over time. This means we want to trace the position of the particule as a function of time, and plot its trajectory. The particle has a position vector $\mathbf{r}(t)$ that changes with time,

$$\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}.$$

The rate at which that position vector changes with respect to time is its velocity vector,

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \frac{\mathrm{d}x}{\mathrm{d}t}\mathbf{i} + \frac{\mathrm{d}y}{\mathrm{d}t}\mathbf{j}.$$

If the particle is immersed in the fluid, it is presumably moving at the same velocity as the fluid, so we should presumably be able to say that

$$\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{v}(x(t), y(t)).$$

The important thing is that the velocity field $\mathbf{v}$ is evaluated *at the current position of the particle*, so at $\mathrm{x} = x(t)$, $y = y(t)$.

But this means that

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v_x(x, y), \qquad \frac{\mathrm{d}y}{\mathrm{d}t} = v_y(x, y).$$

The trajectory of the particle is given by a pair of 'coupled' ordinary differential equations, which can often be solved. The plot of the trajectory in the xy-plane is then called a streamline.[4] To visualize a vector field then means computing and plotting a number of different streamlines, corresonding to different initial positions for the particle.

The relationship between the representation of $\mathbf{v}$ in terms of arrows and in terms of streamlines is that the arrows depicting $\mathbf{v}$ are tangential to the streamlines. The arrows show the direction and speed of travel but do not necessarily make the path taken by an individual particle in the flow obvious; streamlines do the opposite job.

Exercise 2 *Assume the vector field takes the form*

$$\mathbf{v} = x\mathbf{i} - y\mathbf{j}.$$

[4]If you are familiar with the mathematical theory of *dynamical systems*, you should know that the terminology there is slightly different; the curve traced by the particle in the xy-plane is actually called an *orbit*, while the word *trajectory* is used to describe the dependence of the particle position $(x(t), y(t))$ on time.

1. *Set up the differential equations that determine the trajectory of a particle* $x(t)$ *and* $y(t)$ *in the flow.*

2. *Show that these can be solved to give* $x = x_0 \exp(t)$*,* $y = y_0 \exp(-t)$

3. *To plot the trajectory of the particle as a curve in the* xy*-plane,* t *needs to be eliminated so that, for instance,* y *can be expressed as a function of* x*. Show that*
$$y = y_0 x_0 / x$$
unless $x_0 = 0$*.*

4. *Streamlines therefore take the form* $y = C/x$ *with* $C = y_0 x_0 =$ *constant. Sketch streamlines for different values of* C*. Note that there should be streamlines in all parts of the* xy*-plane. Indicate with an arrow the direction in which the particle moves along each streamline.*

Exercise 3 *Sometimes, it makes sense to eliminate* t *imeediately from the calculation of streamlines. We can do this by recognizing that*
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}t} \Big/ \frac{\mathrm{d}x}{\mathrm{d}t}.$$
Therefore we get a single differential equation of the form
$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{v_y(x,y)}{v_x(x,y)}.$$
Apply this approach to find and sketch streamlines for the vector field, again indicating with an arrow the direction in which the particle moves along the streamline.

1. $\mathbf{v} = x\mathbf{i} - y\mathbf{j}$.

2. $\mathbf{v} = -y\mathbf{i} + x\mathbf{j}$.

Evaluating surface integrals: flat surfaces

The abstract notion of a surface integral giving a rate of mass transfer isn't much use if we can't evaluate the integral. We need a practical way of calculating surface integrals. We will develop the relevant tools in two stages. First, we deal with flat surfaces lying in one of the coordinate planes (for simplicity, the xy-plane). Subsequently, we will deal with curved surfaces.

As with volume integrals, precisely how we split the surface S into surface elements δS should not matter, but forcing the δS's to have certain, simple shapes will make life easier. A flat surface in the xy-plane can be covered approximately by small rectangles of edge lengths δx and δy parallel to the x- and y-axes. At the edges of the surface, there may be a slight discrepancy between the area covered by rectangles and the surface S itself, but this will be very small for small δx and δy. The surface elements have size

$$\delta S = \delta x \delta y.$$

The procedure that follows is perfectly analogous to how we constructed volume integrals. We need to compute the sum over

$$\sum_{\delta S} \mathbf{q} \cdot \hat{\mathbf{n}} \delta S. \tag{4}$$

To make the sum over the δS's more definite, we can label each of the rectangles. Let i denote which column a rectangle is in, and j the row, and let (x_i, y_j) be a point associated with the rectangle (e.g., let x_i and y_i be the coordinates of the centre of the rectangle). This allows us to evaluate the normal component of flux $\mathbf{q} \cdot \hat{\mathbf{n}}$.

For a surface S in the xy-plane, the normal $\hat{\mathbf{n}}$ is either $+\mathbf{k}$ or $-\mathbf{k}$. Let us assume that we want to compute the rate of mass transfer from below S to above, so $\hat{\mathbf{n}} = \mathbf{k}$. If $\mathbf{q}$ is written in the form (3), it follows that $\mathbf{q} \cdot \hat{\mathbf{n}} = q_z(x, y, z)$. On the rectangle centered on (x_i, y_j), we have $(x, y, z) = (x_i, y_j, 0)$, The zero appears because, in order to be on the surface S, z cannot just take any value. We have to be in the xy-plane in this particular case, which means $z = 0$.

The sum in (4) becomes

$$\sum_{i,j} q_z(x_i, y_j, 0) \delta x \delta y,$$

When taking a sum over two indices, it does not matter in which order the sum is performed so long as the contribution from each pair of indices (i, j) is included exactly once in the sum. This means that we can sum over j first, and over i second, so

$$\sum \mathbf{q} \cdot \hat{\mathbf{n}} \delta S = \sum_i \left[\sum_j \left. (\mathbf{q} \cdot \hat{\mathbf{n}}) \right|_{x=x_i, y=y_j} \delta y \right] \delta x \tag{5}$$

Taking the sum over j (in square brackets) first means summing over all rectangles in a column first, and then sum over i then amounts to summing up over all the columns next. In the limit of small δy, the sum in square brackets obviously becomes an integral over y, taken at fixed $x = x_i$. Subsequently taking the sum over i then turns into integration with respect to x after the integral with respect to y,

$$\sum_j \left. (\mathbf{q} \cdot \hat{\mathbf{n}}) \right|_{x=x_i, y=y_j} \delta y \to \int \left[\int q_z(x, y, 0) \, \mathrm{d}y \right] \mathrm{d}x \tag{6}$$

As for volume integral, the integral over y is computed with the x-coordinate treated as a constant.

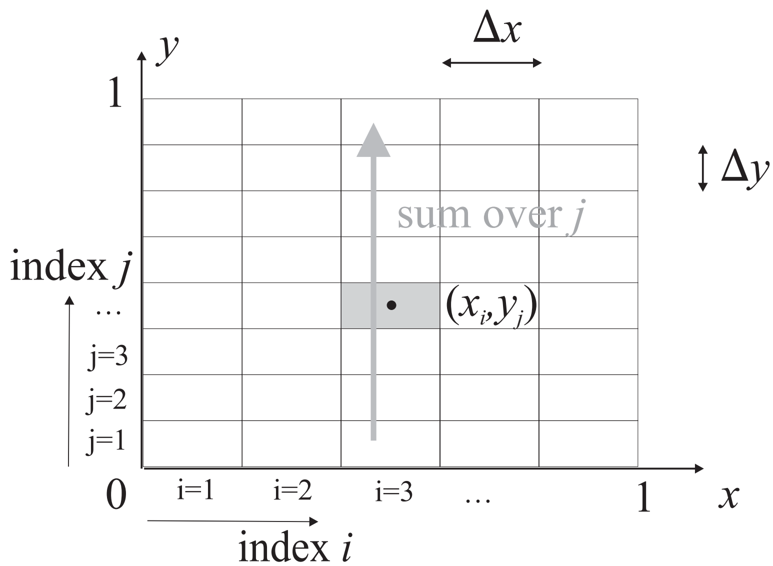

Figure 2: The Riemann sum for a flat, square surface, using rectangular surface elements δS.

We are still mssing limits of integration. Let the boundaries of S be given in the same way as we specified the projection of a volume onto the xy-plane, with a lower boundary at $y = y_{min}(x)$ and an upper boundary at $y = y_{max}(x)$, with x ranging from x_{min} to x_{max}. The integral can then be written in the formula

$$\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \int_{x_{min}}^{x_{max}} \left[\int_{y_{min}(x)}^{y_{max}(x)} q_z(x, y, 0) \, \mathrm{d}y \right] \mathrm{d}x.$$

Example 2 *Let S be the unit square in the xy-plane, and let $\rho = 1$ and*

$$\mathbf{q}(x, y, z) = y\mathbf{i} + x\mathbf{j} + x(1+z)\mathbf{k} \tag{7}$$

We get $x_{min} = 0$, $x_{max} = 1$, $y_{min}(x) = 0$, $y_{max})(x) = 1$.

$$\mathbf{q} \cdot \hat{\mathbf{n}} = (y\mathbf{i} + x\mathbf{j} + x(1+z)\mathbf{k}) \cdot \mathbf{k} = x(1+z).$$

Evaluating this at $z = 0$ gives $\mathbf{q} \cdot \hat{\mathbf{n}} = x$. Hence

$$\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \int_0^1 \left[\int_0^1 x \, \mathrm{d}y \right] \mathrm{d}x = \int_0^1 xy|_{y=0}^{y=1} \, \mathrm{d}x = \int_0^1 x \, \mathrm{d}x = \frac{1}{2}.$$

Example 3 *Consider now the same flux field as given in* (7)*, but assume that S is now the triangle with vertices at* $(0, 0, 0)$*,* $(1, 0, 0)$ *and* $(1, 1, 0)$ *instead. This triangle*

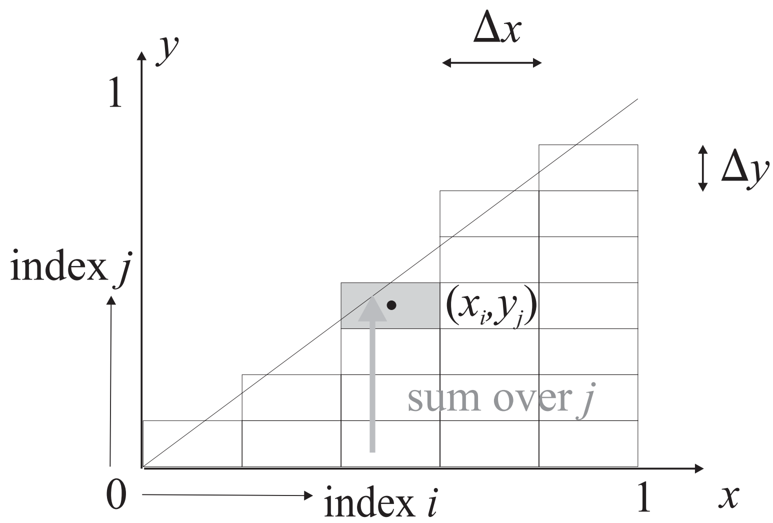

Figure 3: The Riemann sum for a flat, triangular surface, using rectangular surface elements δS. Obviously these do not fit neatly into a triangle. Here we have retained only those δS's whose centres lie in the triangle. As δx *and* δy get smaller, the fit gets better.

is still in the xy-plane, and so we still have $\mathbf{q}\cdot\hat{\mathbf{n}} = x$. *The limits are now different; we have* $y_{max}(x) = x$ *while the other limits remain the same. Therefore*

$$\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \int_0^1 \left[\int_0^x x\,\mathrm{d}y\right]\mathrm{d}x = \int_0^1 xy|_{y=0}^{y=x}\,\mathrm{d}x = \int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}$$

Exercise 4 *Let* S *be a triangle in the xy-plane with vertices at* $(0,0,0)$, $(1,1,0)$, $(-1,1,0)$. *Consider a fluid with density*

$$\rho(x,y,z) = 1 + x^2 + y^2.$$

and velocity field

$$\mathbf{v}(x,y,z) = x\mathbf{i} + y^2\mathbf{j} + (x+1+z)y\mathbf{k}.$$

1. *Sketch the surface* S.
2. *Calculate* $\mathbf{q}\cdot\hat{\mathbf{n}}$ *for the surface* S *and the density and velocity field above.*
3. *Calculate* $\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S$.

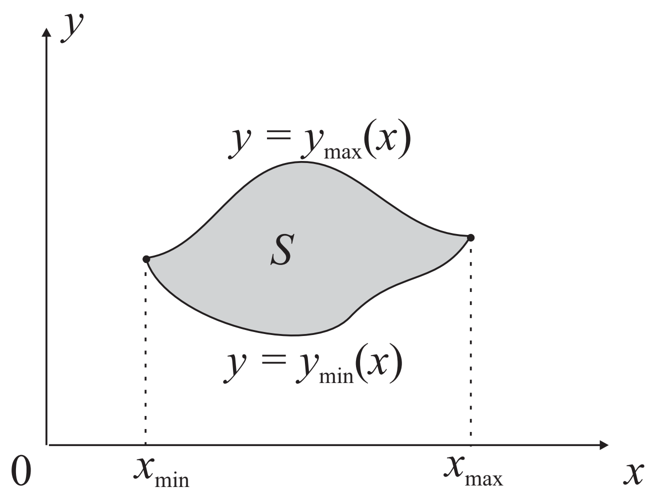

Figure 4: Limits of integration for a more general flat surface.

Curved surfaces

Curved surfaces are more difficult because we cannot simply turn the δS's into little rectangles with side lengths δx and δy, because the δS's do not generally lie in the xy-plane. Moreover, it is not immediately clear how to compute the normal vector $\hat{\mathbf{n}}$. We look at these two issues and how to formulate the surface integral next.

Surface elements and unit normals

The first thing we need to do is specify the shape of the surface. The surface no longer necessarily lies in the xy-plane. Instead, we assume it can be described by its height above the xy-plane, $z = h(x, y)$.

Note 4 *In case a surface curves back on itself, there may be more than one elevation z corresponding to a point on the surface for a given position (x, y) in the xy-plane. In that case, it becomes necessary to split S into multiple parts, each of which corresponds to a unique elevation $z = h(x, y)$ for a given (x, y).*

To get around the problems of described above, we split S into surface elements δS such that the *projection* of each δS onto the xy-plane is a rectangle of side lengths δx and δy (figure 5). What this means is that each δS is a parallelogram, and we can compute the area and even figure out a normal vector for a parallelogram.

To do this, think of the sides of the parallelogram as vectors $\mathbf{a}$ and $\mathbf{b}$ (figure 7). The area of the parallelogram is $|\mathbf{a}||\mathbf{b}| \sin\theta$, where θ is the angle between the two vectors. Recall that the cross product of two vectors is given by

$$\mathbf{a} \times \mathbf{b} = |\mathbf{a}||\mathbf{b}| \sin\theta \mathbf{N}$$

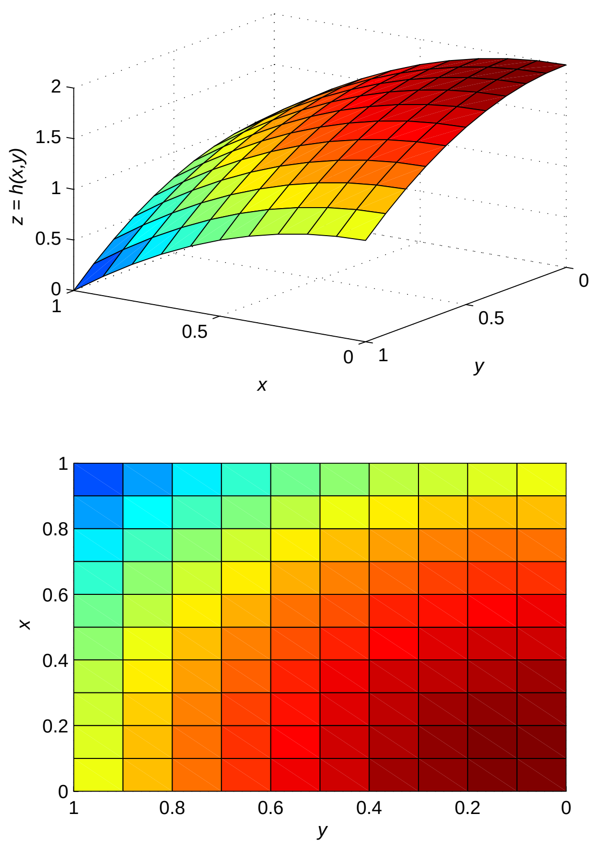

Figure 5: A tiling of the surface $2-x^2-y^2$ with $\delta x = \delta y = 0.1$. Top panel: perspective view, bottom panel: projection onto the xy-plane.

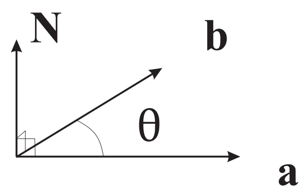

Figure 6: Cross product of two vectors **a** and **b**. **N** is the unit vector perpendicular to **a** and **b**, oriented such that $\mathbf{a}, \mathbf{b}, \mathbf{N}$ form a right-hand triad.

where the unit vector **N** is perpendicular to both **a** and **b**, with the right-hand rule determining the orientation of the vector.

Obviously, if **a** and **b** are the sides of the parallelogram, then **N** is the unit normal $\hat{\mathbf{n}}$ that we are looking for, so

$$\hat{\mathbf{n}} = \frac{\mathbf{a} \times \mathbf{b}}{|\mathbf{a} \times \mathbf{b}|}.$$

Furthermore, the area δS of the parallelogram is the magnitude of $\mathbf{a} \times \mathbf{b}$,

$$\delta S = |\mathbf{a} \times \mathbf{b}|.$$

Therefore we only have to figure out how to write the sides of the each of the surface elements as vectors. To understand this, see figure 7. Take side **a** as the side that projects onto the side of the rectangle in the xy-plane with side length δx. The left-hand end of side **a** is at $(x, y, h(x, y))$, and the right-hand end is at $(x + \delta x, y, h(x + \delta x, y)$, so

$$\begin{aligned}
\mathbf{a} &= [(x + \delta x)\mathbf{i} + y\mathbf{j} + h(x + \delta x, y)\mathbf{k}] - [x\mathbf{i} + y\mathbf{j} + h(x, y)\mathbf{k}] \\
&= \delta x\mathbf{i} + [h(x + \delta x, y) - h(x, y)]\,\mathbf{k}
\end{aligned} \tag{8}$$

We can re-write this slightly by using the definition of a partial derivative:

$$\frac{\partial h}{\partial x} = \lim_{\delta x \to 0} \frac{h(x + \delta x, y) - h(x, y)}{\delta x}$$

From this it follows that

$$h(x + \delta x, y) - h(x, y) \approx \frac{\partial h}{\partial x} \cdot \delta x,$$

so that

$$\mathbf{a} \approx \delta x \left[\mathbf{i} + \frac{\partial h}{\partial x}\mathbf{k}\right],$$

where obviously the quality of the approximation ($\approx$) improves as $\delta x \to 0$.

Switching the roles of x and y and **i** and **j**, we can also find the other side **b** of the parallelogram:

$$\mathbf{b} \approx \delta y \left[\mathbf{j} + \frac{\partial h}{\partial y}\mathbf{k}\right].$$

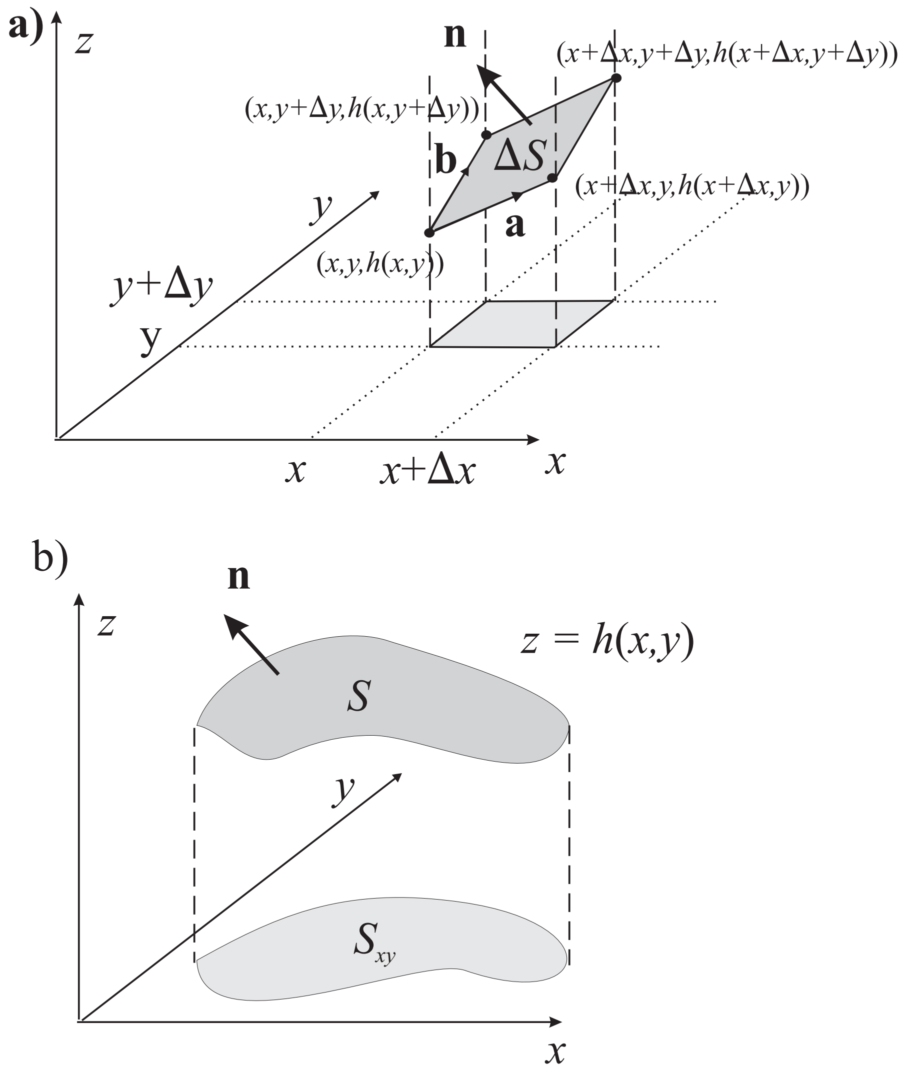

Figure 7: Surface elements δS and unit normal $\hat{\mathbf{n}}$ for a general, curved surface. One can approximate the elevations of the corners as $h(x+\delta x, y) \approx h(x,y) + \partial h/\partial x \cdot \delta x$, $h(x, y+\delta y) \approx h(x,y) + \partial h/\partial y \cdot \delta y$, $h(x+\delta x, y+\delta y) \approx h(x,y) + \partial h/\partial x \cdot \delta x + \partial h/\partial y \cdot \delta y$

Therefore we can calculate

$$
\begin{aligned}
\mathbf{a}\times\mathbf{b} &= \delta x\delta y\left[\mathbf{i}+\frac{\partial h}{\partial x}\mathbf{k}\right]\times\left[\mathbf{j}+\frac{\partial h}{\partial y}\mathbf{k}\right] \\
&= \delta x\delta y\left[\mathbf{i}\times\mathbf{j}+\frac{\partial h}{\partial x}\mathbf{k}\times\mathbf{j}+\frac{\partial h}{\partial y}\mathbf{i}\times\mathbf{k}+\frac{\partial h}{\partial x}\frac{\partial h}{\partial y}\mathbf{k}\times\mathbf{k}\right] \\
&= \delta x\delta y\left[\mathbf{k}-\frac{\partial h}{\partial x}\mathbf{i}-\frac{\partial h}{\partial y}\mathbf{j}\right]
\end{aligned}
$$

From this, it follows that

$$
\begin{aligned}
\delta S &= |\mathbf{a}\times\mathbf{b}| \\
&= \delta x\delta y\left|\mathbf{k}-\frac{\partial h}{\partial x}\mathbf{i}-\frac{\partial h}{\partial y}\mathbf{j}\right| \\
&= \delta x\delta y\sqrt{1+\left(\frac{\partial h}{\partial x}\right)^2+\left(\frac{\partial h}{\partial y}\right)^2}
\end{aligned}
$$

and

$$
\begin{aligned}
\hat{\mathbf{n}} &= \frac{\mathbf{a}\times\mathbf{b}}{|\mathbf{a}\times\mathbf{b}|} \\
&= \frac{\mathbf{k}-\frac{\partial h}{\partial x}\mathbf{i}-\frac{\partial h}{\partial y}\mathbf{j}}{\sqrt{1+\left(\frac{\partial h}{\partial x}\right)^2+\left(\frac{\partial h}{\partial y}\right)^2}}
\end{aligned}
$$

Exercise 5 *So far, we have taken it on trust that the surface element δS really is a parallelogram, and computed vectors describing only* two *of its sides. Compute the vectors describing the third and fourth sides of the element and show that, when δx and δy are small, those additional two sides take the same form as the vectors* **a** *and* **b** *that we have already computed. This shows that we really do have a parallelogram.*

Example 4 *Let S be the triangle with vertices $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ (see figure 8). Calculate $\hat{\mathbf{n}}$ and δS.*

The first thing we need to do is to write the surface in the form $z = h(x, y)$. We know that the triangle is part of a plane, and hence we expect that

$$z = ax + by + c. \tag{9}$$

describes the surface for some set of coefficients a, b and c. How do we find them? We know that the plane must pass through the points $(x, y, z) = (1, 0, 0)$, $(0, 1, 0)$

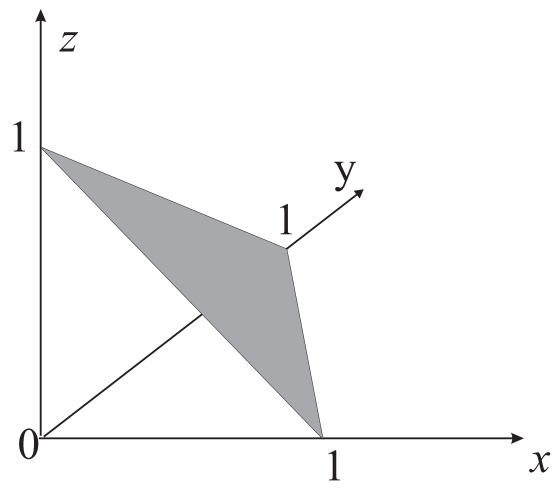

Figure 8: The surface S in example 4.

and $(0,0,1)$. *Hence these coordinates must satisfy the equation (9). Substituting the coordinates into (9), we get, for each of these points in turn,*

$$\begin{aligned} 0 &= a + c \\ 0 &= b + c \\ 1 &= c \end{aligned}$$

Hence $c = 1$, $a = b = -1$ *and* $z = 1 - x - y$. *Therefore*

$$h(x, y) = 1 - x - y.$$

and

$$\frac{\partial h}{\partial x} = -1, \qquad \frac{\partial h}{\partial y} = -1, \qquad \sqrt{1 + \left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2} = \sqrt{3}$$

From the formulae above,

$$\begin{aligned} \hat{\mathbf{n}} &= \frac{\mathbf{k} + \mathbf{i} + \mathbf{j}}{\sqrt{3}}, \\ \delta S &= \sqrt{3}\delta x \delta y. \end{aligned} \tag{10}$$

Rate of mass transfer $\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S$: evaluating $\mathbf{q}$ on the surface and determining limits of integration

Hence we can write the normal component of flux as

$$\mathbf{q} \cdot \hat{\mathbf{n}} = \frac{\mathbf{q} \cdot \left[\mathbf{k} - \frac{\partial h}{\partial x}\mathbf{i} - \frac{\partial h}{\partial y}\mathbf{j}\right]}{\sqrt{1 + \left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}.$$

Multiplying by δS, we get

$$\mathbf{q}\cdot\hat{\mathbf{n}}\delta S = \frac{\mathbf{q}\cdot\left[\mathbf{k}-\frac{\partial h}{\partial x}\mathbf{i}-\frac{\partial h}{\partial y}\mathbf{j}\right]}{\sqrt{1+\left(\frac{\partial h}{\partial x}\right)^2+\left(\frac{\partial h}{\partial y}\right)^2}}\cdot\sqrt{1+\left(\frac{\partial h}{\partial x}\right)^2+\left(\frac{\partial h}{\partial y}\right)^2}\delta x\delta y$$

$$= \mathbf{q}\cdot\left[\mathbf{k}-\frac{\partial h}{\partial x}\mathbf{i}-\frac{\partial h}{\partial y}\mathbf{j}\right]\delta x\delta y \tag{11}$$

and so

$$\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \int\int \mathbf{q}\cdot\left[\mathbf{k}-\frac{\partial h}{\partial x}\mathbf{i}-\frac{\partial h}{\partial y}\mathbf{j}\right]\mathrm{d}y\,\mathrm{d}x.$$

This last expression shows that we get a reasonably simplified expression for the integral. In particular, if we are given the flux in component form, $\mathbf{q} = q_x\mathbf{i}+q_y\mathbf{j}+q_z\mathbf{k}$, then we end up with the simpler-looking double integral

$$\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \int\int\left[q_z - q_x\frac{\partial h}{\partial x} - q_y\frac{\partial h}{\partial y}\right]\mathrm{d}y\,\mathrm{d}x. \tag{12}$$

There is an important point here. Remember that q_x, q_y and q_z are functions of x, y and z. When we sum over the terms $\mathbf{q}\cdot\hat{\mathbf{n}}\delta S$, the normal component of flux $\mathbf{q}\cdot\hat{\mathbf{n}}$ must be evaluated on the surface. This means again that z is again not arbitrary, but must be $z = h(x,y)$.The integrand in (12) therefore becomes

$$\int\int\left[q_z(x,y,h(x,y)) - q_x(x,y,h(x,y))\frac{\partial h}{\partial x} - q_y(x,y,h(x,y))\frac{\partial h}{\partial y}\right]\mathrm{d}y\,\mathrm{d}x. \tag{13}$$

We still need limits of integration. Recall that, as in the case of a flat surface, we are simply adding over surface elements δS. Equivalently, we are adding over the projections of these δS's onto the xy-plane. We chose these projections to be rectangles with sides parallel to the x- and y-axes, and so we are back to adding all the rectangles in one column (the y-integral) and then adding all the columns of rectangles (the x-integral).

In practice, what this means for the limits of integration is tha they are given by the boundaries of the projection S_{xy} onto the xy-plane. If S_{xy} takes the form indicated for S in figure 4, then the limits of integration are the same as for the flat, planaer surface case we discussed before. The integral over the cureved surface finally takes the form

$$\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \int_{x_{\min}}^{x_{\max}}\int_{y_{\min}(x)}^{y_{\max}(x)}\left[q_z(x,y,h(x,y)) - q_x(x,y,h(x,y))\frac{\partial h}{\partial x} - q_y(x,y,h(x,y))\frac{\partial h}{\partial y}\right]\mathrm{d}y\,\mathrm{d}x.$$

We put this formula into practice next.

Example 5 *Let S be the surface $h = 2 - x^2 - y^2$ that lies above the triangle in the xy-plane with vertices $(0,0)$, $(1,0)$ and $(0,1)$, and let $\mathbf{q} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Calculate $\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S$.*

From the information given, we have

$$\frac{\partial h}{\partial x} = -2x, \qquad \frac{\partial h}{\partial y} = -2y$$

while on the surface S, flux is

$$\mathbf{q} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} = x\mathbf{i} + y\mathbf{j} + (2 - x^2 - y^2)\mathbf{k}.$$

so $q_x(x, y, h(x, y)) = x$, $q_y(x, y, h(x, y)) = y$ and $q_z(x, y, h(x, y)) = 2 - x^2 - y^2$. Hence the integrand in (13) is

$$-x\cdot(-2x) - y\cdot(-2y) + 2 - x^2 - y^2 = 2 + x^2 + y^2.$$

Now for the limits of integration. The projection of S onto the xy-plane is the triangle with vertices $(0,0)$, $(1,0)$ and $(0,1)$, so the lower boundary is $y_{\min} = 0$ while the upper boundary is $y_{\max} = 1 - x$, and the left-hand end-point is at $x_{\min} = 0$, while the right-hand edge is $x_{\max} = 1$ (draw the triangle to see this). So

$$\begin{aligned}
\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S &= \int_0^1 \left[\int_0^{1-x} 2 + x^2 + y^2\,\mathrm{d}y\right]\mathrm{d}x \\
&= \int_0^1 \left[2y + x^2 y + y^3/3\right]_0^{1-x}\,\mathrm{d}x \\
&= \int_0^1 2(1-x) + x^2(1-x) + (1-x)^3/3\,\mathrm{d}x \\
&= \int_0^1 2 - 2x + x^2 - x^3 + (1-x)^3/3\,\mathrm{d}x \\
&= \left[2x - x^2 + x^3/3 - x^4/4 - (1-x)^4/12\right]_0^1 \\
&= \frac{7}{6}
\end{aligned}$$

Exercise 6 *Let $\mathbf{q} = y\mathbf{i} - x\mathbf{j} + z\mathbf{k}$, and $h(x, y) = \cos(x)\cos(y)$. Calculate $\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S$, where S is the part of the surface $z = h(x, y)$ that lies above the square in the xy-plane for which $0 < x < \pi/2$, $0 < y < \pi/2$.*

Exercise 7 *What do you think the physical interpretation of $\int_S 1\,\mathrm{d}S$ is? Confirm your answer by evaluating this integral for two triangles, one with vertices $(0,0,0)$, $(0,1,0)$, $(1,0,0)$, and the other with vertices $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, and comparing your answer with the areas of these triangles computed directly from the formula $1/2\times$base$\times$height.*

For $\mathbf{q}\cdot\mathbf{n} = q_n =$ constant, evaluate $\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S$ if A is the area of surface S. Why is this the answer you expect from the definition of flux?

The orientation of $\hat{\mathbf{n}}$ and surface integrals over closed surfaces

Above, we computed the normal vector $\hat{\mathbf{n}}$ as

$$\hat{\mathbf{n}}_+ = \frac{\mathbf{k} - \frac{\partial h}{\partial x}\mathbf{i} - \frac{\partial h}{\partial y}\mathbf{j}}{\sqrt{1 + \left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}}, \tag{14}$$

where we have introduced the subscript '+' to indicate that the vector points upwards. Equally possible would have been the unit vector with the opposite direction

$$\hat{\mathbf{n}}_- = -\hat{\mathbf{n}}_+, \tag{15}$$

which instead points downwards.

As discussed at the start of these notes, using $\hat{\mathbf{n}}_-$ instead of $\hat{\mathbf{n}}_+$ in the rate of mass transfer calculation $\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S$ would simply lead to a change of sign in the integral. This has the simple physical meaning we discussed before: either choice of normal vector points from one side of the surface to the other, and the integral $\int_S \mathbf{q} \cdot \hat{\mathbf{n}}$ represents mass transfer in the same direction. For instance, $\mathbf{q} \cdot \hat{\mathbf{n}}_+$ is the normal component of flux that points from below the surface $z = h(x, y)$ to above, and correspondingly $\int_S \mathbf{q} \cdot \hat{\mathbf{n}}_+ \, \mathrm{d}S$ is the rate of mass transfer from below the surface to above. Similarly $\int_S \mathbf{q} \cdot \hat{\mathbf{n}}_- \, \mathrm{d}S$ is the rate of mass transfer from above the surface to below. The choice of which unit normal must be used is therefore given by the direction of mass transfer that we are interested in.

It often happens that we want to calculate the rate of mass transfer across a closed surface, as this this tells us the rate at which the mass contained within that surface changes. A closed surface S is one which fully encloses a volume V. To compute the rate at which mass leaves the volume V, we have to calculate $\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S$, where $\hat{\mathbf{n}}$ is the *outward-pointing* unit normal.

Note 5 *For a closed surface, convention states that* $\hat{\mathbf{n}}$ *always points* out *of the surface.*

A further complication generally arises with closed surfaces: they cannot be represented as a single surface of the form $z = h(x, y)$. Instead, a closed surface must be broken into several pieces, for each of which one of the formulae we have developed can be applied. Of course, there is nothing special about the z-axis, and it may turn out that one of the pieces of the closed surface has to be represented in the form $y = h(x, z)$ or $x = h(y, z)$, in which case the roles of x, y and z in the formulae above must be interchanged accordingly. We illustrate this with an example.

Example 6 *Let V be the tetrahedron with vertices $(0, 0, 0)$, $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ (figure 9), and let S be its surface, with outward-pointing unit normal $\hat{\mathbf{n}}$. Let $\mathbf{q} = (x - y - z)\mathbf{i} + (y + x)\mathbf{j} + (x + z)\mathbf{k}$. Compute $\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S$.*

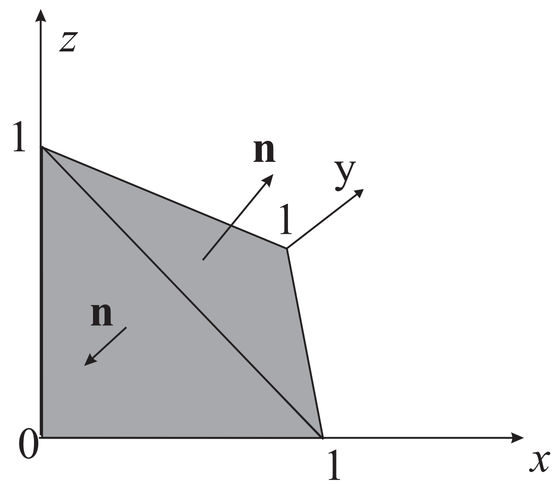

Figure 9: The volume V in example 6.

The obvious first step is to split S into the faces of the tetrahedron. The slanted face is the same as in example (4) above. Label this S_1. In addition, there are two vertical faces, one in the xz-plane (facing the viewer in figure 9 and the other in the yz-plane (obscured in figure 9). Label these S_2 and S_3, respectively. Lastly, there is the base of the tetrahedron, in the xy-plane, which we label S_4. These faces are illustrated in figure 10.

The surface integral $\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S$ can be found by finding the corresponding surface integral over each face $S_1,\ldots,S_4$ and summing:

$$\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \sum_{i=1}^{4}\int_{S_i}\mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S.$$

We consider each face in turn.

For S_1, we have a surface of the form $z = h(x,y) = 1 - x - y$, and $\partial h/\partial x = \partial h/\partial y = -1$. Moreover, we know that the outward-pointing unit normal is also upward-pointing, so

$$\hat{\mathbf{n}} = \frac{\mathbf{k} - \frac{\partial h}{\partial x}\mathbf{i} - \frac{\partial h}{\partial y}\mathbf{j}}{\sqrt{1 + \left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}} = \frac{\mathbf{k} + \mathbf{i} + \mathbf{j}}{\sqrt{3}}.$$

On the surface, with $z = 1 - x - y$, we have $\mathbf{q} = (2x-1)\mathbf{i} + (y+x)\mathbf{j} + (1-y)\mathbf{k}$, and so

$$\mathbf{q}\cdot\hat{\mathbf{n}} = \frac{(2x-1) + (y+x) + (1-y)}{\sqrt{3}} = \frac{3x}{\sqrt{3}} = \sqrt{3}x.$$

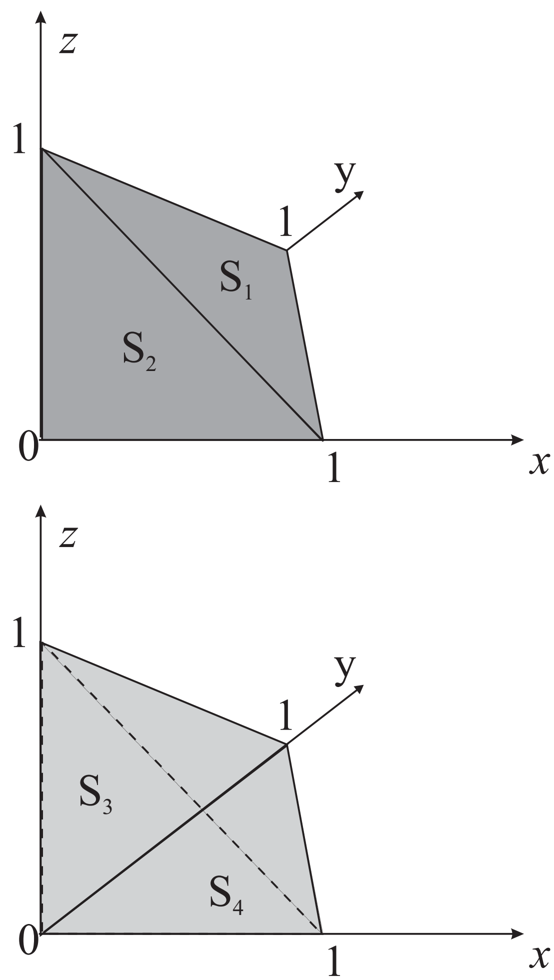

Figure 10: The faces making up S in example 6.

Moreover, a surface element δS can be written as

$$\delta S = \sqrt{1 + \left(\frac{\partial h}{\partial x}\right)^2 + \left(\frac{\partial h}{\partial y}\right)^2}\delta x \delta y = \sqrt{3}\delta x \delta y.$$

Hence $\int_{S_1} \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = \int\int \sqrt{3}x\sqrt{3}\,\mathrm{d}y\,\mathrm{d}x$. To get the limits of integration, we have to look at the projection of S_1 onto the xy-plane. This is simply the triangle with vertices at $(0,0,0)$, $(1,0,0)$ and $(0,1,0)$ (in other words, it is the face S_4 in figure 10). Its upper and lower boundaries can be written as $y_{\min}(x) = 0$ and $y_{\max}(x) = 1-x$, respectively, while its left-hand end is at $x_{\min} = 0$ and its right-hand end is at $x_{\max} = 1$. Hence

$$\begin{aligned}
\int_{S_1} \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S &= \int_0^1 \int_0^{1-x} 3x\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_0^1 3xy\big|_{y=0}^{y=1-x}\,\mathrm{d}x \\
&= \int_0^1 3x - 3x^2\,\mathrm{d}x \\
&= \left[\frac{3}{2}x^2 - x^3\right]_{x=0}^{x=1} \\
&= \frac{1}{2}.
\end{aligned}$$

S_2 clearly cannot be written in the form $z = h(x,y)$. However, as it lies in the xz-plane, it is straightforward to see that $\delta S = \delta x \delta z$. The outward-pointing unit normal points in the negative y-direction, and so $\hat{\mathbf{n}} = -\mathbf{j}$. Moreover, we have $y = 0$ on S_2, so $\mathbf{q} = (x-z)\mathbf{i} + x\mathbf{j} + (x+z)\mathbf{k}$. Therefore

$$\mathbf{q}\cdot\hat{\mathbf{n}} = -x.$$

The boundaries of the triangle are at $x_{\min} = 0$ and $x_{\max} = 1$, $z_{\min}(x) = 0$ and $z_{\max}(x) = 1-x$, so

$$\begin{aligned}
\int_{S_2} \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S &= \int_0^1 \int_0^{1-x} -x\,\mathrm{d}z\,\mathrm{d}x \\
&= \int_0^1 -xz\big|_0^{1-x}\,\mathrm{d}x \\
&= \int_0^1 -x(1-x)\,\mathrm{d}x \\
&= \left[-\frac{x^2}{2} + \frac{x^3}{3}\right]_0^1 \\
&= -\frac{1}{6}
\end{aligned}$$

Similarly, S_3 lies in the yz-plane, with $\delta S = \delta y \delta z$ and boundaries at $z_{\min} = 0$, $z_{\max} = 1 - y$, $y_{\min} = 0$ and $y_{\max} = 1$. Moreover, the outward-pointing unit normal is $\hat{\mathbf{n}} = -\mathbf{i}$ and with $x = 0$ on S_3, we have

$$\mathbf{q} \cdot \hat{\mathbf{n}} = y + z.$$

Hence

$$\begin{aligned}
\int_{S_3} \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S &= \int_0^1 \int_0^{1-y} y + z \, \mathrm{d}z \, \mathrm{d}y \\
&= \int_0^1 \left[yz + \frac{z^2}{2} \right]_0^{1-y} \mathrm{d}y \\
&= \int_0^1 y(1-y) + \frac{(1-y)^2}{2} \, \mathrm{d}y \\
&= \int_0^1 \frac{1-y^2}{2} \, \mathrm{d}y \\
&= \frac{1}{3}
\end{aligned}$$

Lastly, S_4 is in the xy-plane, with $\delta S = \delta x \delta y$ and outward-pointing unit normal $-\mathbf{k}$. On S_4, $z = 0$, and so

$$\mathbf{q} \cdot \hat{\mathbf{n}} = -x,$$

while the boundaries of S_4 are $x_{\min} = 0$, $x_{\max} = 1$, $y_{\min}(x) = 0$, $y_{\max}(x) = 1 - x$. Hence

$$\begin{aligned}
\int_{S_3} \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S &= \int_0^1 \int_0^{1-x} -x \, \mathrm{d}y \, \mathrm{d}x \\
&= \int_0^1 -x(1-x) \, \mathrm{d}x \\
&= \left[-\frac{x^2}{2} + \frac{x^3}{3} \right]_0^1 \\
&= -\frac{1}{6}
\end{aligned}$$

Summing all the integrals, we get

$$\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \sum_{i=1}^{4} \int_{S_i} \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = \frac{1}{2} - \frac{1}{6} + \frac{1}{3} - \frac{1}{6} = \frac{1}{2}.$$

Exercise 8 *Repeat the previous example but with V the tetrahedron with vertices $(0,0,0)$, $(1,0,0)$, $(0,1,0)$ and $(1,0,1)$, and S its surface. Show that you still get $\int_S \mathbf{q} \cdot \hat{\mathbf{n}} \, \mathrm{d}S = 1/2$.*

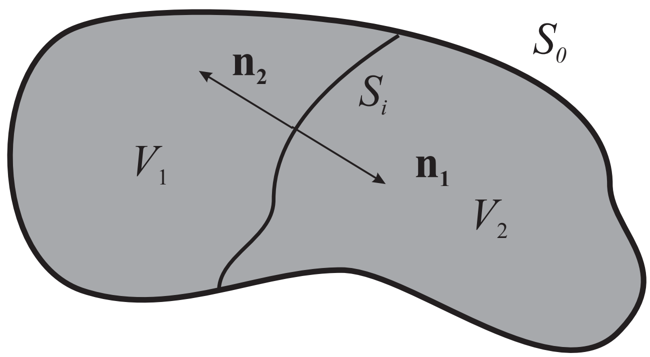

Figure 11: Mass exchange between two volumes V_1 and V_2.

Conservation laws

Now we have a way of computing the mass contained in a general volume V, and the rate at which mass flows through a surface. Can we connect the two? If mass is conserved, then the rate at which mass in the volume increases must equal the rate of mass flow in, or equivalently, the negative of the rate at which mass flows out through the surface of the volume. Therefore, if S is the surface of V and $\hat{\mathbf{n}}$ is an outward-pointing unit normal to S, then

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho \,\mathrm{d}V = -\int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S,$$

or in other words,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho \,\mathrm{d}V + \int_S \mathbf{q}\cdot\hat{\mathbf{n}}\,\mathrm{d}S = 0. \tag{16}$$

Importantly, the calculation above is not restricted to any particular volume V, but must hold for *any* volume and its surface S.

At this point, you might still ask, how does (16) imply mass conservation? The mass contained in V, $\int_V \rho\,\mathrm{d}V$, manifestly need not stay constant. The point is not that the mass of in any particular volume needs to stay the same, just as conservation of momentum does not mean that the momentum of a billiard ball needs to stay the same in a collision: instead the total momentum of all billiard balls involved in the collision needs to stay the constant. The equivalent of this here is the observation that if mass flows out of one volume V, it must flow into a neighbouring volume.

Suppose you had two adjoining volumes V_1 and V_2 that together make up a larger volume V_0, with surface S_1 and S_2, with outward-pointing unit normals $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$. Let S_i be the surface that S_1 and S_2 have in common (i.e., along which V_1 and V_2 meet) and let S_0 be the outer surface of V_0 (figure 11). Suppose that there is no flow out of the composite volume V_0, so $\mathbf{q}\cdot\hat{\mathbf{n}} = 0$ on S_0, but that there can be flow between

V_1 and V_2. Then the flow out of V_1 is then

$$-\frac{\mathrm{d}}{\mathrm{d}t}\int_{V_1} \rho \, \mathrm{d}V = \int_{S_1} \mathbf{q}\cdot\hat{\mathbf{n}}_1 \, \mathrm{d}S = \int_{S_i} \mathbf{q}\cdot\hat{\mathbf{n}}_1 \, \mathrm{d}S,$$

as there is no flow through the other parts of S_1. Similarly, the flow out of V_2 is

$$-\frac{\mathrm{d}}{\mathrm{d}t}\int_{V_1} \rho \, \mathrm{d}V = \int_{S_2} \mathbf{q}\cdot\hat{\mathbf{n}}_2 \, \mathrm{d}S = \int_{S_i} \mathbf{q}\cdot\hat{\mathbf{n}}_2 \, \mathrm{d}S.$$

As the total mass in V is

$$\int_V \rho \, \mathrm{d}V = \int_{V_1} \mathrm{d}V + \int_{V_2} \rho \, \mathrm{d}V,$$

we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho \, \mathrm{d}V &= \frac{\mathrm{d}}{\mathrm{d}t}\int_{V_1} \rho \, \mathrm{d}V + \frac{\mathrm{d}}{\mathrm{d}t}\int_{V_2} \rho \, \mathrm{d}V \\
&= -\int_{S_i} \mathbf{q}\cdot\hat{\mathbf{n}}_1 \, \mathrm{d}S - \int_{S_i} \mathbf{q}\cdot\hat{\mathbf{n}}_2 \, \mathrm{d}S \\
&= -\int_{S_i} \mathbf{q}\cdot\left[\hat{\mathbf{n}}_1 + \hat{\mathbf{n}}_2\right] \, \mathrm{d}S
\end{aligned}$$

But $\hat{\mathbf{n}}_1$ and $\hat{\mathbf{n}}_2$ are equal and opposite in size, so $\hat{\mathbf{n}}_1 + \hat{\mathbf{n}}_2 = \mathbf{0}$ and $\int_{S_i} \mathbf{q}\cdot[\hat{\mathbf{n}}_1 + \hat{\mathbf{n}}_2] = 0$. The flow of mass out of V_1 and V_2 is equal and opposite, and the total amount of mass is conserved:

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho \, \mathrm{d}V = 0.$$