

## Chapter 2

# Differential calculus in many dimensions

Descriptions of physical systems can have many state variables. The state variables don't always assume arbitrary independent values. For instance, suppose there are three state variables  $x$ ,  $y$  and  $z$  and physically admissible triples live on some surface, as depicted in Figure 2.1a. A surface is a *graph* if one state variable, say  $z$ , is a function of the other two,  $x$  and  $y$ . In this case,  $x$  and  $y$  are in the role of *independent* variables, and  $z$  in the role of *dependent*. Typically, a surface is composed of “graphs glued together”, and at most points on a surface, one can choose two of  $x$ ,  $y$  or  $z$  as independent variables, and the remaining one dependent.

For the shaded patch of surface in Figure 2.1a, we represent  $z = z(x, y)$  as a function of  $x$  and  $y$ . In the “magnified” picture (Figure 1.2b) the curve  $ab$  is the intersection of the surface with a plane  $y = \text{constant}$ . On this curve,  $z$  is a function of  $x$ , and its derivative with respect to  $x$ , denoted  $z_x$  (or more traditionally  $\frac{\partial z}{\partial x}$ ) is called the *partial derivative of  $z$  with respect to  $x$* . Project the curve  $ab$  onto the  $x, z$  plane. You get the graph of a function of one variable ( $x$ ) and  $z_x$  represents the slopes of its tangent lines. Similar meaning of  $z_y$ : In Figure 2.1b,  $cd$  is intersection of surface with a plane  $x = \text{constant}$ . Project  $cd$  onto the  $yz$  plane, and  $z_y$  represents the slopes of tangent lines in the  $y, z$  plane. Higher derivatives have clear meanings.  $z_{xx}$  means “differentiate with respect to  $x$ , and then differentiate  $z_x$  with respect to  $x$ ”.  $z_{xy}$  means “differentiate with respect to  $y$ , and then differentiate with respect to  $x$ ”. This seems backward. Its inherited from the

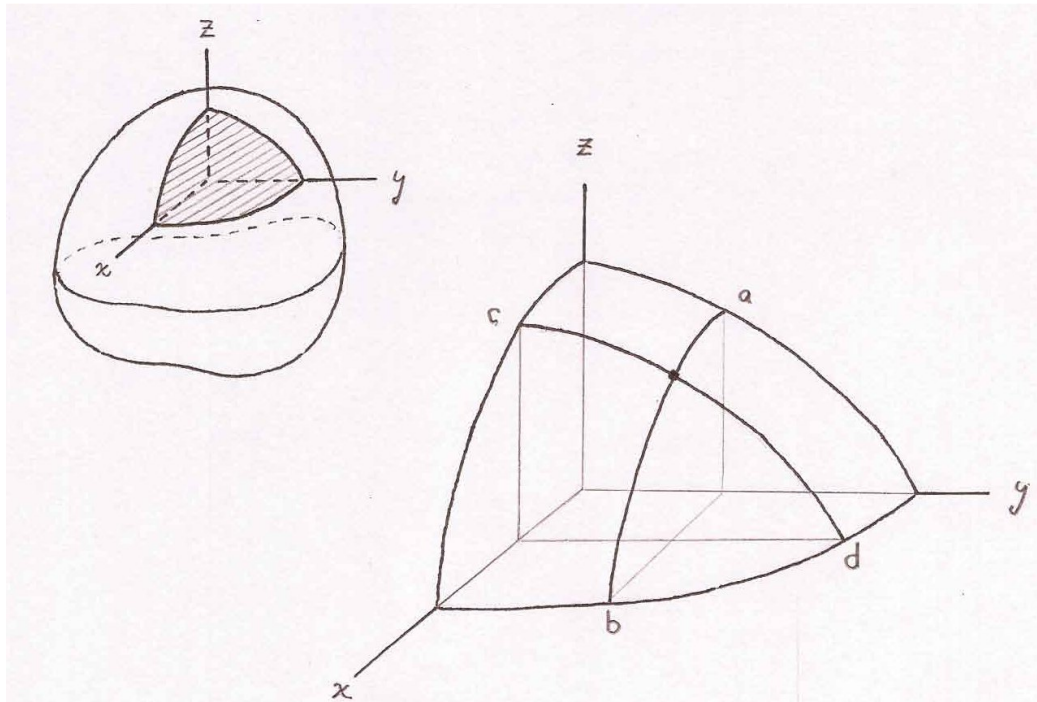


Figure 2.1

“traditional” notation  $\frac{\partial^2 z}{\partial x \partial y} := \frac{\partial}{\partial x} \left( \frac{\partial z}{\partial y} \right)$ . Most of the time, the order of  $x$  and  $y$  differentiations in “mixed” partial derivatives does not matter: If  $z_{xy}$  and  $z_{yx}$  are both continuous, then they are equal:  $z_{xy} = z_{yx}$ .

You’ve seen most of this before in calculus. What we’ll concentrate on here is the *interchangeability of roles between independent and dependent variables*. Certain sciences, mechanics and thermodynamics in particular, are notorious for “always changing their minds about the roles of variables, independent or dependent”. This gives rise to notations that you don’t see in math books: For instance, suppose you decide that  $x$  is the dependent variable, and  $y, z$  independent, and you want to differentiate with respect to  $y$ . A notation you commonly see in thermodynamics is  $\left( \frac{\partial x}{\partial y} \right)_z$ . The parentheses and subscript  $z$  are a reminder that “the  $y$  differentiation happens with  $z$  fixed”. Of course, you might argue that the  $( )_z$  is redundant, and it is, so long as  $x, y$ , and  $z$  are the *only* variables. But suppose there are four,  $x, y, z$ , and  $s$ , only two of which are independent. If we write  $z_x$ , which of  $(z_x)_y$

or  $(z_x)_s$  do we mean?

Here are some elementary exercises: The surface is the “saddle”

$$z = x^2 - y^2 \quad (2.1)$$

depicted in Figure 2.2. Suppose we decide that  $x$  is the dependent variable.

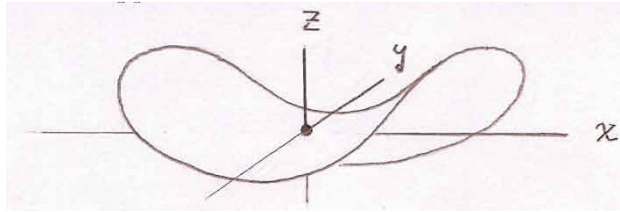


Figure 2.2

If you look at lines parallel to the  $x$  axis, you’ll see that lines with  $z < -y^2$  *don’t* intersect the surface at all, and lines with  $z > -y^2$  intersect it *twice*. This is reflected in the math: “Solving” (2.1) for  $x$  gives

$$x = \pm\sqrt{z + y^2} \quad (2.2)$$

but only for  $z > -y^2$ . Given (2.2), we now calculate

$$x_y = \pm \frac{y}{\sqrt{z + y^2}}. \quad (2.3)$$

To visualize why there are these two values of  $x_y$ , recall that the  $y$ -differentiation happens with  $z$  *fixed*, so we should look at  $z$ -level curves of (2.1), which are the hyperbolas in Figure 2.3. For given  $y > 0$ , the *positive* value of  $x_y$  is the slope of the tangent line at  $a$ , and the *negative* value, the slope of the tangent line at  $a'$ .

It is common to transform a given set of independent variables into others. For instance, suppose we start with  $x$  and  $y$  as independent variables, but we replace  $y$  by  $s := x + y$  or  $t := x - y$ . If the independent variables are  $x, s$ , we have  $z = (x + y)(x - y) = s(2x - s)$ , and

$$\left(\frac{\partial z}{\partial x}\right)_s = 2s.$$

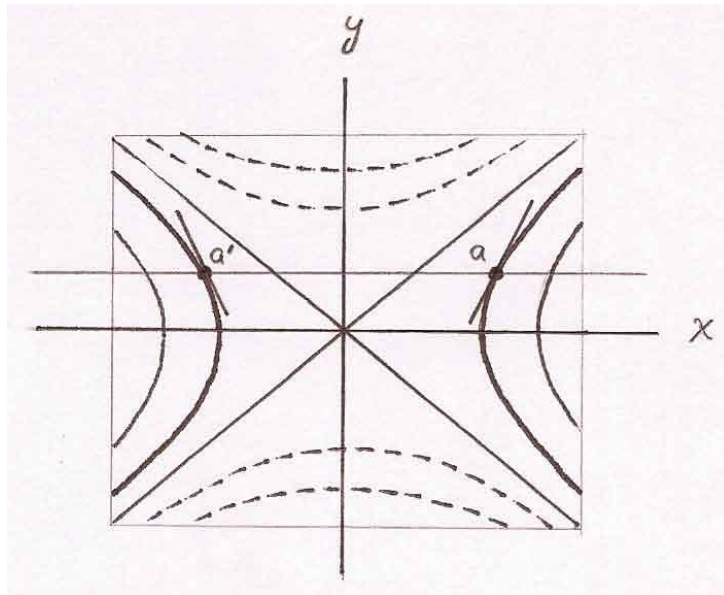


Figure 2.3

If we use  $x, t$ , we have  $z = (2x - t)t$  and

$$\left(\frac{\partial z}{\partial x}\right)_t = 2t.$$

Notice that the  $( )_s, ( )_t$  notations come into their own.

This last example is a preamble to the famous question from gas thermodynamics: What is the pressure as a function of volume if the temperature is a uniform constant? How does the pressure-volume relation change if there is no heat transfer? The basic state variables of a gas are  $v :=$  volume/molecule,  $p :=$  pressure,  $e :=$  energy/molecule,  $\tau :=$  temperature, and  $\sigma :=$  entropy. Hence, we have a five-dimensional state space with coordinates  $v, p, e, \tau, \sigma$ . Physically realizable *equilibria* distinguish a two-dimensional surface in this five-dimensional space. That is, *two* independent variables and the remaining three, dependent. For instance, a simple mechanical model of an ideal gas (non-interacting point particles) called kinetic theory leads to

$$pv = \tau, \quad e = \frac{3}{2}\tau. \quad (2.4)$$

In addition, the entropy  $\sigma$  may be represented as a function of  $v$  and  $\tau$ . The notion of entropy is notoriously subtle. For the moment it is sufficient

to recall that change in entropy indicates heat transfer to or from the gas, and “heat transfer” means “exchange of energy with surroundings which is not mechanical work”. In an exercise, you’ll investigate the dependence of entropy upon  $v$  and  $\tau$  for an ideal gas. In summary, we have a description of an ideal gas at equilibrium, with three relations between the five state variables. Hence, two independent and three dependent variables.

We return to the original question, about the pressure-volume relation at constant temperature, or subject to no heat transfer. It boils down to comparing  $\left(\frac{\partial p}{\partial v}\right)_\tau$  and  $\left(\frac{\partial p}{\partial v}\right)_\sigma$ .

For isothermal ( $\tau = \text{constant}$ ) expansion done in a *reversible* manner (slowly, so you are always near equilibrium) we have from (2.4),

$$p = \frac{\tau}{v},$$

and then

$$\left(\frac{\partial p}{\partial v}\right)_\tau = -\frac{\tau}{v^2} = -\frac{p}{v}. \quad (2.5)$$

Here is the “energy budget” for this expansion: As  $v$  increases from  $v_1$  to  $v_2$ , the gas does *work* (per molecule)

$$W = \int_{v_1}^{v_2} p dv = \tau \int_{v_1}^{v_2} \frac{dv}{v} = \tau \log \frac{v_2}{v_1}$$

and the required energy is absorbed from the surroundings (called a “heat bath”, and “heat”  $\tau \log \frac{v_2}{v_1}$  is absorbed from the bath).

Now suppose the expansion happens, but with *no* heat transfer from the surroundings. The gas now has constant entropy  $\sigma$  and we’re investigating  $\left(\frac{\partial p}{\partial v}\right)_\sigma$ . Although we haven’t presented a formula for  $\sigma$ , we can still figure out  $\left(\frac{\partial p}{\partial v}\right)_\sigma$  from a modified energy budget: This time the work done by the gas is paid for by a drop in its energy  $e$  per molecule, so

$$\frac{de}{dt} = -(\text{rate of work}) = -p \frac{dv}{dt}. \quad (2.6)$$

Here,  $t$  is time and the time dependences are assumed to be sufficiently slow that this process is also “reversible”. Substituting  $e = \frac{3}{2}\tau$  and  $p = \frac{\tau}{v}$  from (2.4) into (2.6), we have

$$\frac{3}{2} \frac{d\tau}{dt} = -\frac{\tau}{v} \frac{dv}{dt}$$

or

$$\frac{1}{\tau} \frac{d\tau}{dt} + \frac{2}{3} \frac{1}{v} \frac{dv}{dt} = 0.$$

It follows that

$$\tau v^{\frac{2}{3}} = c(\sigma)$$

where  $c(\sigma)$  is a time independent constant that presumably depends on the similarly constant value of  $\sigma$ . Hence,  $\tau = \frac{c(\sigma)}{v^{\frac{2}{3}}}$  and

$$p = p(v, \sigma) = \left( \frac{c(\sigma)}{v^{\frac{2}{3}}} \right) \frac{1}{v} = \frac{c(\sigma)}{v^{\frac{5}{3}}}. \quad (2.7)$$

Finally, we compute

$$\left( \frac{\partial p}{\partial v} \right)_{\sigma} = -\frac{5}{3} \frac{c(\sigma)}{v^{\frac{8}{3}}}, \quad (2.8)$$

or eliminating  $c(\sigma)$  by means of (2.7),

$$\left( \frac{\partial p}{\partial v} \right)_{\sigma} = -\frac{5}{3} \frac{p}{v}. \quad (2.9)$$

Comparing (2.5), (2.9), we see that

$$\left( \frac{\partial p}{\partial v} \right)_{\sigma} = \frac{5}{3} \left( \frac{\partial p}{\partial v} \right)_{\tau}. \quad (2.10)$$

Figure 2.4 is a pictorial summary of (2.10).

### Multidimensional power series

Let  $z(x, y)$  be *analytic* in the sense that  $z(x, y)$  has all partial derivatives in some region about a given point  $(x, y) = (a, b)$ . We want to construct “Taylor polynomials in two dimensions” which give good asymptotic approximations to  $z(x, y)$  as  $(x, y) \rightarrow (a, b)$ . First, in the one-variable Taylor series (1.48), set  $x = a + h$  to obtain

$$f(a + h) = \sum_{m=0}^{\infty} \frac{f^{(m)}(a)}{m!} h^m. \quad (2.11)$$

Next, look at  $z(a + h, y)$  with  $y$  fixed. Following (2.11), we can make the Taylor series in powers of  $h$ ,

$$z(a + h, y) = \sum_{m=0}^{\infty} \frac{1}{m!} ((\partial_x)^m z)(a, y) h^m. \quad (2.12)$$

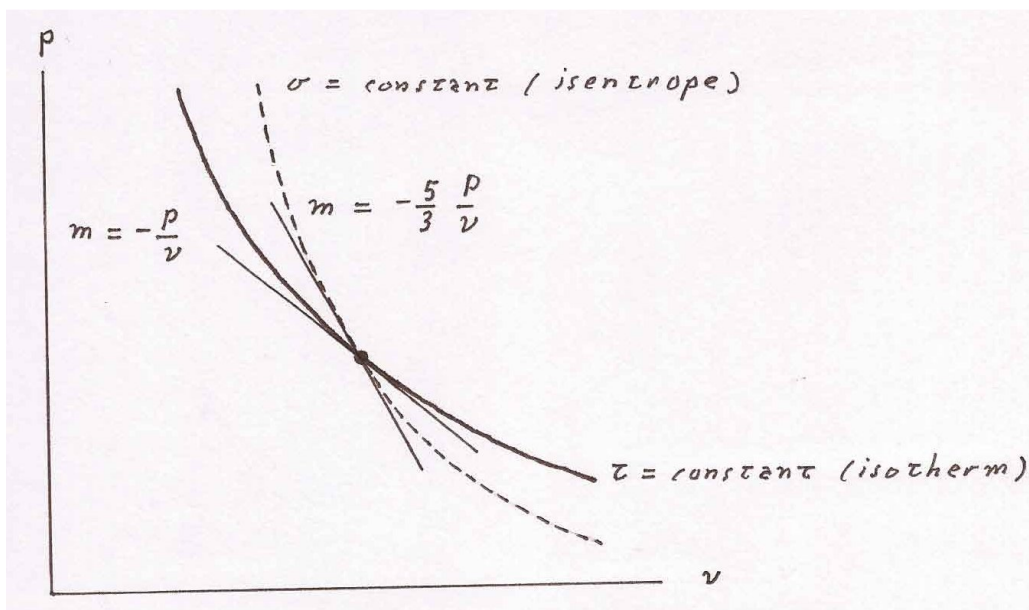


Figure 2.4

Here, the notation  $((\partial_x)^m z)(a, y)$  means: “ $m$   $x$ -partial derivatives of  $z$  evaluated at  $(a, y)$ ”. Similarly,

$$z(a, b + k) = \sum_{n=0}^{\infty} \frac{1}{n!} ((\partial_y)^n z)(a, b) k^n.$$

We can do exactly the same to  $((\partial_x)^m z)(a, b + k)$ , so

$$((\partial_x)^m z)(a, b + k) = \sum_{n=0}^{\infty} \frac{1}{n!} ((\partial_x)^m (\partial_y)^n z)(a, b) k^n. \quad (2.13)$$

Combining (2.12), (2.13), we present “the whole catastrophe”

$$z(a + h, b + k) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \frac{1}{m!} \frac{1}{n!} ((\partial_x)^m (\partial_y)^n z)(a, b) h^m k^n. \quad (2.14)$$

There is a rearrangement of the summations that helps this terrible formula: For given  $N \geq 0$ , sum over  $m, n$  so  $m + n = N$ . Then sum over  $N$ . This

leads to

$$\sum_{N=0}^{\infty} \sum_{m=0}^N \frac{1}{N!} \frac{N!}{m!(N-m)!} ((\partial_x)^m (\partial_y)^{N-m} z)(a, b) h^m k^{N-m}. \quad (2.15)$$

This stinks of the binomial formula,

$$(h + k)^N = \sum_{m=0}^N \binom{N}{m} h^m k^{N-m}.$$

Fine, but how do we slip in the derivatives  $((\partial_x)^m (\partial_y)^{N-m} z)(a, b)$ ? By means of the *directional derivative operator*

$$(h\partial_x + k\partial_y)z(x, y) := hz_x(x, y) + kz_y(x, y). \quad (2.16)$$

Formally, we have

$$((h\partial_x + k\partial_y)^N z)(a, b) = \sum_{m=0}^N \binom{N}{m} ((\partial_x)^m (\partial_y)^{N-m} z)(a, b) h^m k^{N-m}$$

and then (2.14) reduces to

$$z(a + h, b + k) = \sum_{N=0}^{\infty} \frac{1}{N!} ((h\partial_x + k\partial_y)^N z)(a, b). \quad (2.17)$$

This still looks terrible. But at least its a “terrible that is easy to remember”. In addition, good news: Most of the time, all we want are the  $N = 0, 1, 2$  terms. Again, the reason is that “low order Taylor polynomials make good approximations as  $(h, k) \rightarrow (0, 0)$ ”.

### Differentials and tangent planes

Very often we need to approximate changes in dependent variables due to small changes in the independent. Let

$$\Delta z := z(a + dx, b + dy) - z(a, b) \quad (2.18)$$

be the change in  $z(x, y)$  when  $(x, y)$  changes from  $(a, b)$  to  $(a + dx, b + dy)$ . Here, the changes  $dx$  and  $dy$  of  $x$  and  $y$  are called *differentials* of  $x$  and  $y$ . (No,



they are *not* “infinitesimals”, so we don’t have to answer to Bishop Berkeley’s critique of Newton, that he traffics in the “ghosts of vanished quantities”.) The approximation to  $\Delta z$  based on the  $N = 0, 1$  terms of the two-variable Taylor series (2.17) with  $h = dx$  and  $k = dy$  is

$$dz = z_x(a, b)dx + z_y(a, b)dy, \quad (2.19)$$

called the *differential of  $z$* . The difference  $\Delta z - dz$  is the sum of terms in (2.17) with  $N \geq 2$ . The  $N = 2$  term has components proportional to  $(dx)^2$ ,  $dx dy$  and  $(dy)^2$ . Let  $dr := \sqrt{(dx)^2 + (dy)^2}$  be the length of the displacement  $(dx, dy)$ . We have  $(dx)^2$ ,  $|dx dy|$  and  $(dy)^2$  all less than  $(dr)^2$ , and the  $N = 2$  term in (2.17) has absolute value bounded above by a constant times  $(dr)^2$ . Similarly, the  $N > 2$  term is bounded by a constant times  $(dr)^N$ . As  $dr \rightarrow 0$ ,  $(dr)^N$  for  $N > 2$  becomes negligible relative to  $(dr)^2$ . It can be shown that

$$\Delta z - dz = O((dr)^2),$$

so

$$\Delta z = z_x(a, b)dx + z_y(a, b)dy + O((dr)^2). \quad (2.20)$$

Figure 2.5 is a geometric visualization of  $\Delta z$  and how it differs from  $dz$ . Notice that  $(a + dx, b + dy, z + dz)$  lies on the *tangent plane* of the graph of  $z(x, y)$  at  $(x, y) = (a, b)$ . In this sense,  $dz$  is called the *tangent plane approximation to  $\Delta z$* .

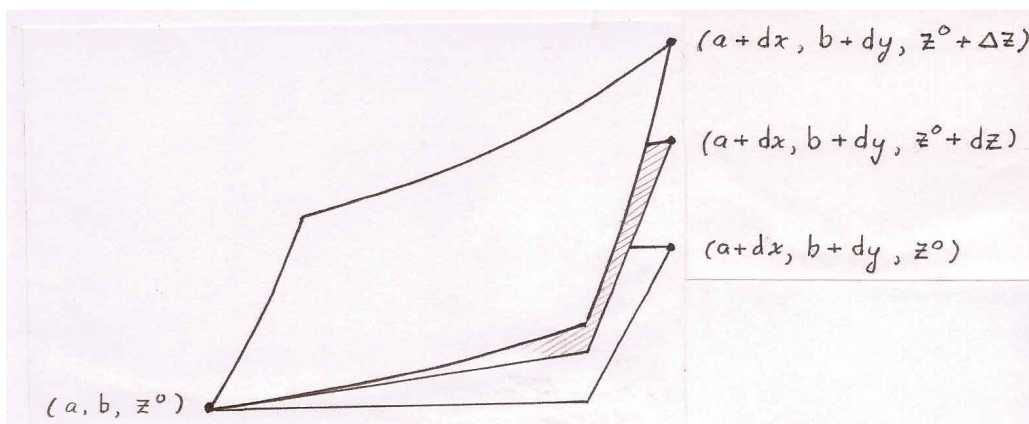


Figure 2.5

We examine two applications of (2.20). (1) The draw weight  $f$  of an archery bow is proportional to  $\omega\tau^3$ , where  $\omega$  and  $\tau$  are the width and thickness

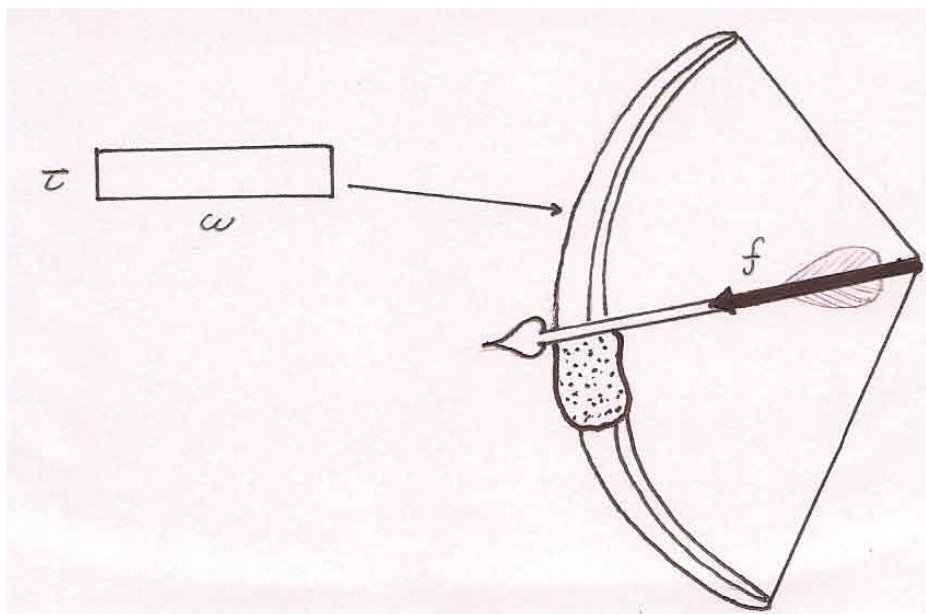


Figure 2.6

of its limbs, say at the middle. Suppose we decrease the thickness  $\tau$  by 2%. By what percentage do we need to increase the width, to keep the same draw weight. By what percentage does the mass of limit per unit length change? From

$$f = \text{constant} \times \omega \tau^3 \quad (2.21)$$

we have

$$df = f_\omega d\omega + f_\tau d\tau = \text{constant} \times (\tau^3 d\omega + 3\omega \tau^2 d\tau) \quad (2.22)$$

and dividing (2.22) by (2.21),

$$\frac{df}{f} = \frac{d\omega}{\omega} + 3\frac{d\tau}{\tau}. \quad (2.23)$$

If  $df = 0$ , and  $\frac{d\tau}{\tau} = -.02$  as given, we get  $\frac{d\omega}{\omega} = +.06$  or 6%. The mass per unit length  $\rho$  is proportional to the cross sectional area  $\omega\tau$ , so

$$\frac{d\rho}{\rho} = \frac{d\omega}{\omega} + \frac{d\tau}{\tau} = .06 - .02 = .04,$$

and the mass per unit length increases by 4%.

(2) The electric potential in the  $x, y$  plane due to uniform charge  $\sigma$  per unit length along the  $z$  axis is

$$u(x, y) = -\sigma \log \sqrt{x^2 + y^2}. \quad (2.24)$$

We have a *dipole* in the  $x, y$  plane, consisting of a charge  $-Q$  at  $(x, y) \neq (0, 0)$  and a charge  $+Q$  at  $(x + dx, y + dy)$ . The electric potential energy of the

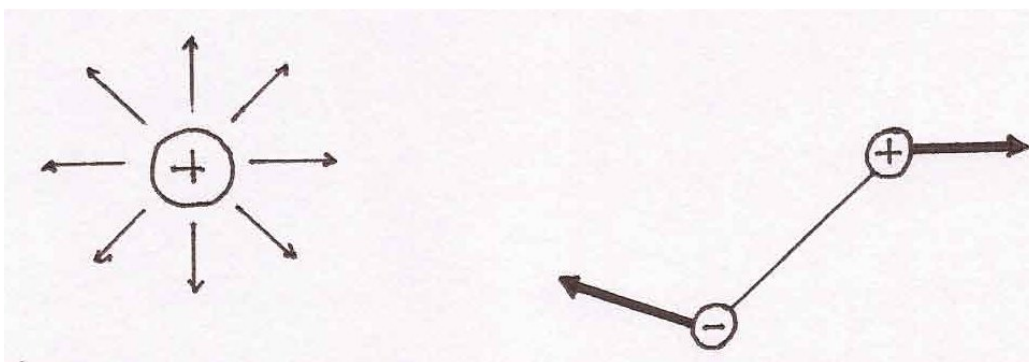


Figure 2.7

dipole is

$$U := Q\sigma u(x + dx, y + dy) - Q\sigma u(x, y) = du + O((dr)^2), \quad (2.25)$$

where

$$du := u_x(x, y)dx + u_y(x, y)dy = -Q\sigma \left( \frac{xdx + ydy}{x^2 + y^2} \right). \quad (2.26)$$

Some vector notation is informative: Let  $\mathbf{x} := (x, y)$  denote displacement from the origin, and  $d\mathbf{x} := (dx, dy)$  the displacement from the negative to the positive charge. Then (2.26) can be written as

$$du = -Q\sigma \frac{\mathbf{x} \cdot d\mathbf{x}}{r^2}$$

where  $r := \sqrt{x^2 + y^2}$  is distance of the negative charge from the origin. Introducing the radial unit vector  $\hat{\mathbf{r}} = \frac{\mathbf{x}}{r}$ , and the *dipole moment*  $\mathbf{p} := Qd\mathbf{x}$ , we have

$$du = -\sigma \frac{\hat{\mathbf{r}} \cdot \mathbf{p}}{r},$$

and finally

$$U = -Q\sigma \frac{\hat{\mathbf{r}} \cdot \mathbf{P}}{r} + O((dr)^2). \quad (2.27)$$

We can guess what happens to the dipole in Figure 2.7: The dark arrows are the forces on the charges due to the line charge along the  $z$  axis. These forces will cause  $\mathbf{p}$  to rotate until it points in the  $+\hat{\mathbf{r}}$  direction, and then the dipole as a whole gets pulled to the origin.

### Multivariable chain rule

Given analytic  $z(x, y)$  and a parametric curve  $(x(t), y(t))$  in the  $x, y$  plane, define  $Z(t)$  as the *restriction* of  $z(x, y)$  to the curve. That is

$$Z(t) := z(x(t), y(t)). \quad (2.28)$$

If  $z = z(x)$  independent of  $y$ , (2.28) reduces to

$$Z(t) = z(x(t)),$$

and by the chain rule for one variable functions, we have

$$\frac{dZ}{dt}(t) = \frac{dz}{dx}(x(t)) \frac{dx}{dt}(t).$$

What is the generalized formula for  $\frac{dZ}{dt}$  when  $z(x, y)$  depends on both  $x$  and  $y$ ?

Let us compute the change  $\Delta Z$  in  $Z$  due to a change  $dt$  in  $t$ , two ways: First, assuming that  $Z(t)$  is given, we have

$$\Delta Z = \frac{dZ}{dt}(t)dt + O((dt)^2). \quad (2.29)$$

Alternatively,

$$\begin{aligned} \Delta Z &= z(x(t+dt), y(t+dt)) - z(x(t), y(t)) \\ &= z(x + \Delta x, y + \Delta y) - z(x, y). \end{aligned}$$

In the last line  $x, y$  denote values at  $t$ , and  $\Delta x := x(t+dt) - x(t)$ , and similarly for  $\Delta y$ . By the differential approximation formula (2.20) (with  $\Delta x$  and  $\Delta y$  in roles of  $dx, dy$ ) we have

$$\Delta Z = z_x(x, y)\Delta x + z_y(x, y)\Delta y + O((\Delta r)^2), \quad (2.30)$$

where  $\Delta r := \sqrt{(\Delta x)^2 + (\Delta y)^2}$ . Next we evoke

$$\Delta x = \frac{dx}{dt} dt + O((dt)^2), \quad (2.31)$$

and similarly for  $\Delta y$ . Substituting these  $\Delta x$  and  $\Delta y$  into (2.30), we obtain

$$\Delta Z = \left\{ z_x(x, y) \frac{dx}{dt} + z_y(x, y) \frac{dy}{dt} \right\} dt + O((dt)^2). \quad (2.32)$$

Notice that the  $O((dt)^2)$  in (2.32) comes from the  $O((dt)^2)$  in (2.31) and  $O((\Delta r)^2)$  in (2.30). Comparing (2.29), (2.32), we see that

$$\frac{dZ}{dt} = z_x(x, y) \frac{dx}{dt} + z_y(x, y) \frac{dy}{dt}.$$

The generalization to  $z$  depending on more independent variables is clear: If

$$Z(t) := z(x_1(t), x_2(t), \dots, x_n(t)) \quad (2.33)$$

we have

$$\frac{dZ}{dt} = z_{x_1} \frac{dx_1}{dt} + \dots + z_{x_n} \frac{dx_n}{dt}.$$

Because of the subscripts on the  $x$ 's, maybe

$$\frac{dZ}{dt} = \frac{\partial z}{\partial x_1} \frac{dx_1}{dt} + \dots + \frac{\partial z}{\partial x_n} \frac{dx_n}{dt} \quad (2.34)$$

looks better. Finally, if the  $x$ 's themselves are functions of several variables  $t_1, t_2, \dots, t_m$ , then for each  $j = 1, \dots, m$  we have

$$\frac{\partial Z}{\partial t_j} = \frac{\partial z}{\partial x_1} \frac{\partial x_1}{\partial t_j} + \dots + \frac{\partial z}{\partial x_n} \frac{\partial x_n}{\partial t_j}. \quad (2.35)$$

## Convective derivative

In continuum mechanics, the evolution of a material body is quantified by a *flow map*. The flow map is a transformation of spatial positions at time  $t = 0$  to positions at times  $t \neq 0$ . If the flow map sends an initial region  $R(0)$  at time  $t = 0$  to the region  $R(t)$  at time  $t$ , then  $R(0)$  and  $R(t)$  “consist of exactly the same material stuff”. A time sequence  $R(t)$  of regions

always containing the same stuff is called a *material region*. Flow maps are a mathematical way of describing a flowing stream or the shaking of a fat man's belly. In the limit of material regions shrinking to moving points, we get "particle trajectories"  $\mathbf{x} = \mathbf{x}(t)$ . At any time  $t$ , the set of velocities  $\dot{\mathbf{x}}(t)$  defines a vector field  $\mathbf{u}(x, t)$  in space, called the *velocity field* of the medium. Each particle trajectory  $\mathbf{x}(t)$  satisfies the ODE

$$\dot{\mathbf{x}} = \mathbf{u}(\mathbf{x}, t). \quad (2.36)$$

For instance, in two-dimensional space,  $\mathbf{u}$  has  $x$  and  $y$  components which we denote  $u(x, y, t)$  and  $v(x, y, t)$ , and the ODE (2.36) takes component form,

$$\frac{dx}{dt} = u(x, y, t), \quad \frac{dy}{dt} = v(x, y, t). \quad (2.37)$$

Figure 2.8 is a pictorial summary of flow maps and velocity fields.

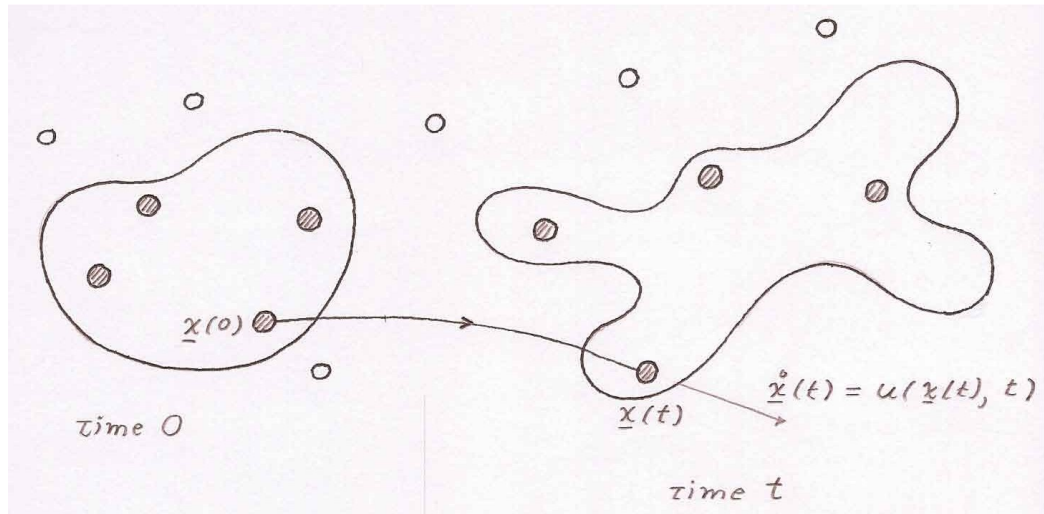


Figure 2.8

The most essential mathematical idea in continuum mechanics is quantifying the time rate of change of a state variable  $c(\mathbf{x}, t)$  along particle trajectories. Here, the multidimensional chain rule is indispensable: Along a particle trajectory  $\mathbf{x} = \mathbf{x}(t)$ , the value of  $c$  as a function of time is

$$C(t) := c(\mathbf{x}(t), t). \quad (2.38)$$

Lets consider two space dimensions and write in place of (2.38),

$$C(t) = c(x(t), y(t), t). \quad (2.39)$$

Applying the multidimensional chain rule (2.34) with  $x_1(t) = x(t)$ ,  $x_2(t) = y(t)$ ,  $x_3(t) = t$ , we have

$$\frac{dC}{dt} = c_x \frac{dx}{dt} + c_y \frac{dy}{dt} + c_t \frac{dt}{dt}. \quad (2.40)$$

Since the  $x$  and  $y$  velocities  $\frac{dx}{dt}$  and  $\frac{dy}{dt}$  satisfy (2.37), we have

$$\frac{dC}{dt} = (c_t + uc_x + vc_y)(x(t), y(t), t). \quad (2.41)$$

The quantity  $c_t + uc_x + vc_y$  is called the *convective derivative* of  $c(x, y, t)$ . We use our understanding of the convective derivative to explain how

### Shear increases gradients

Let  $c(x, y, t)$  be the concentration of dye in flowing water. Water is essentially incompressible, so any little blob of water maintains the same volume, and the concentration of dye inside it remains constant in time as well. Hence,  $C(t)$  in (2.39) is time independent, and by (2.40), the convective derivative of  $c$  vanishes, so we have the PDE

$$c_t + uc_x + vc_y = 0. \quad (2.42)$$

If  $u$  and  $v$  are given, this PDE governs the time evolution of  $c(x, y, t)$ . Let's consider the special case of a *shear flow* with  $u = -\omega y$ ,  $v \equiv 0$ , where  $\omega$  is a constant with units of  $1 \div \text{time}$ . The PDE (2.42) reduces to

$$c_t - \omega y c_x = 0. \quad (2.43)$$

Figure 2.9 depicts the shear flow and its effect on a blob of dye. Particle trajectories have  $y \equiv \text{constant}$ , independent of time, and  $x = -\omega y t$ , modulo an additive constant. Intuitively, we understand that an initially circularly symmetric dye blob elongates in a “streak” of ever increasing length that gradually aligns itself parallel to the  $x$  axis.

Lets see what happens to the gradient  $\nabla c = (c_x, c_y)$  seen along a particle path. That is, we examine the time evolutions of  $F(t) := c_x(-\omega y t, y, t)$  and

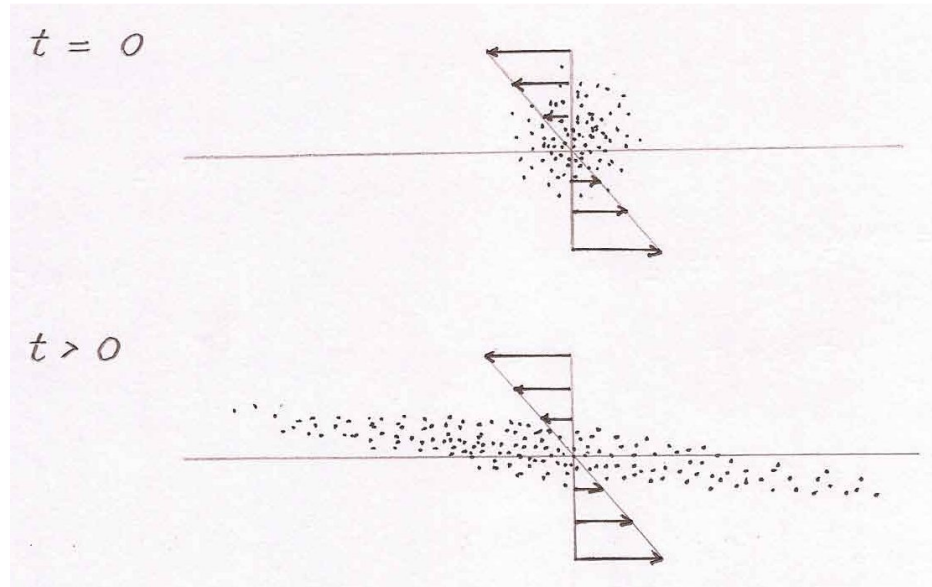


Figure 2.9

$G(t) := c_y(-\omega yt, y, t)$ , where  $y$  is any constant. The time derivatives of  $F(t)$  and  $G(t)$  are convective derivatives of  $c_x$  and  $c_y$ ,

$$\frac{dF}{dt} = (c_x)_t - \omega y (c_x)_x, \quad (2.44)$$

$$\frac{dG}{dt} = (c_y)_t - \omega y (c_y)_x. \quad (2.45)$$

In (2.44), (2.45), the derivatives of  $c$  are evaluated at  $x = -\omega yt$ ,  $y = \text{constant}$ . By  $x$ -differentiation of the PDE (2.43) we see that the convective derivative of  $c_x$  vanishes and then (2.44) implies  $F \equiv \text{constant}$ , independent of  $t$ . By  $y$ -differentiation of (2.43), we find

$$(c_y)_t - \omega y (c_y)_x - \omega c_x = 0$$

and then (2.45) reduces to

$$\frac{dG}{dt} = \omega c_x(-\omega yt, y) = \omega F.$$

Since  $F \equiv \text{constant}$ , we have

$$G(t) = \omega Ft,$$



modulo an additive constant. In summary, the  $x$  gradient  $c_x$  is constant in time along a particle path, and the  $y$  gradient  $c_y$  grows linearly in time. Evidently, the blob in Figure 2.9 is becoming *narrow* in the  $y$  direction. That makes sense: As it is elongated in the  $x$ -direction, it must *compress* in the  $y$  direction to preserve its area.