Chapter 2
Differential calculus

This chapter is written for those people who have the courage to approach the mathematics of general relativity without being familiar with differential calculus. The use of this fabulous creation by Newton and Leibniz is essential and omnipresent on our way to Einstein's field equations.

2.1 Differentiation

In Fig. 2.1 we have drawn Cartesian coordinates and a curve. Five lines lead from a point P on the curve to points Q_1, Q_2, Q_3, Q_4, Q_5, also on the curve. If we continue plotting lines like that, the Q's approach P indefinitely. They are said to approach the tangent line at P. The lines PQ_1, PQ_2, PQ_3, PQ_4, and PQ_5 have different slopes in relation to the x-axis. The slope of the curve at a point P is defined as the slope of the tangent line at P. Knowing the slope of a curve at any point, and the value of the function at one point, we can plot the curve (Fig. 2.2). If the the tangent lines are close enough the curve 'plots itself'. Could we find simple expressions for the slope of those lines, that is, for the slope of the curve itself? The slope k of a line is quantitatively expressed by what we shall call the 'slope quotient' defined as 'increment in y-direction divided by increment in x-direction' (see Fig. 2.3).

$$k = \frac{y_2 - y_1}{x_2 - x_1}. \tag{2.1}$$

In Fig. 2.2 the slope quotient at P is $1/5 = 0.20$. The slopes of the series of lines PQ_1, PQ_2, PQ_3, ... in Fig. 2.2 are expressed by the same sort of quotient as that of the line pictured in Fig. 2.3. Their slope quotients approach that of the tangent line, as Q_n approaches P.
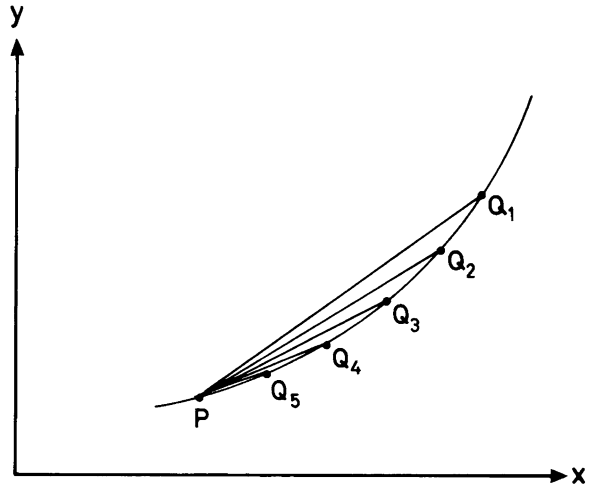
Fig. 2.1 A curve

Fig. 2.2 A curve with
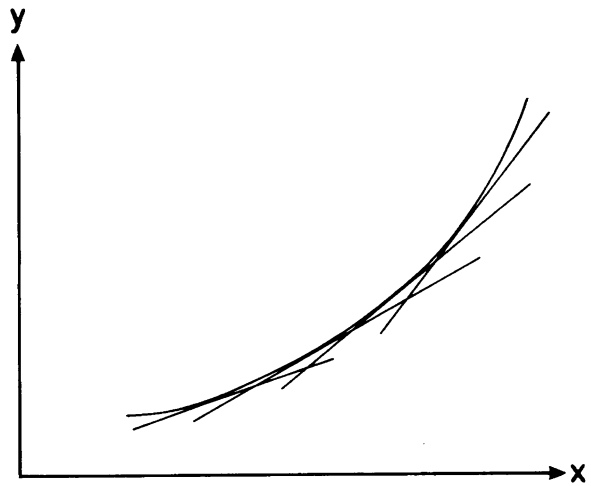tangent lines

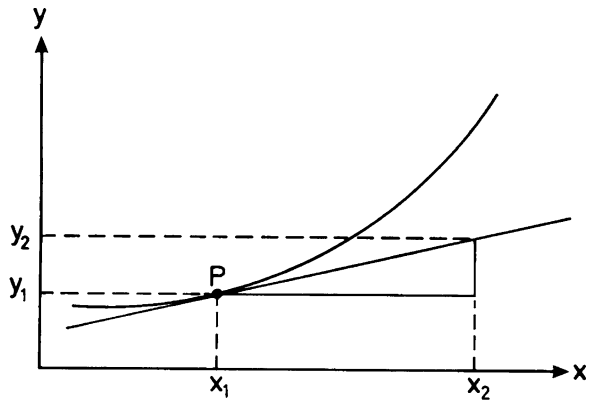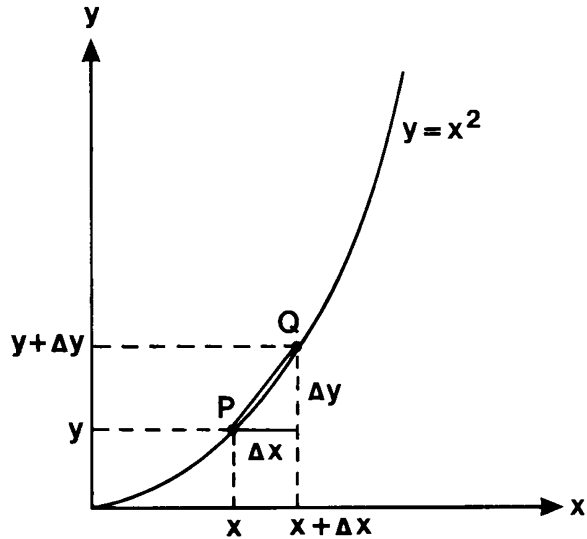**Fig. 2.3** The slope of the
tangent line

Fig. 2.4 The parabola
$y = x^2$

The slope quotient of the curve on Fig. 2.2 increases continuously, from less than 1/5 to a fairly big number. That of Fig. 2.3 seems to start from zero and remain there for a while, like a straight line parallel to the x-axis.

A word of caution: not all curves have a tangent at every point. If the curve has a sharp corner at a point it does not have a tangent at that point. Also, if the curve has a vertical step (the vertical interval is not reckoned as part of the curve), then it is disconnected for a certain value of x, say $x = x_1$. This is called a discontinuity. If the curve is the graph of a function, the function is said to have a discontinuity at $x = x_1$. A curve does not have a tangent at a discontinuity. We shall assume that the functions we need to consider, are such that their graphs have one definite tangent line at every point, i.e. that the functions are continuous and their graphs are without sharp corners or discontinuous steps. If you inspect Fig. 2.4 you will see that a new notation has been introduced. The figure suggests that we move along a curve, called a parabola, with equation $y = x^2$, a short distance from P to Q. If P is an arbitrary point on the curve, we denote the coordinates of P by x and y, and Q has coordinates $x + \Delta x$ and $y + \Delta y$, where Δx ('delta x') is a small increment added to x, and Δy a corresponding (i.e. related to Δx and the steepness of the curve, see Fig. 2.4) small increment of y.

In the new notation we may now express in a general way the slope quotient of the curve making use of Eq. (2.1), one page 21:

$$\frac{(y + \Delta y) - y}{(x + \Delta x) - x} = \frac{\Delta y}{\Delta x}. \tag{2.2}$$

However close the points P and Q on a curve may be, there is a definite quotient $\Delta y / \Delta x$. When points on the fairly simple curves (graphs of continuous and singularity free functions) we shall study are brought closer and closer, the quotient approaches a definite limit, which represents the slope of the curve at the point P. There is a symbol for the approach towards this 'limit' which some of us find quite elegant:

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}. \tag{2.3}$$

Let us now consider the curve in Fig. 2.4 as the graph of a function $y = f(x)$. The limit (2.3) changes when x varies except when the curve, in our wide sense of the term, is a straight line. Therefore the limit is evidently a function of x, and is called *the derivative* of $f(x)$. Note that $f(x)$ and the derivative of $f(x)$ are two different functions. The process of finding the derivative is not called derivation but 'differentiation'. As could be expected there are many symbols expressing this crucial notion. The most intuitively powerful is perhaps dy/dx. Since $y = f(x)$ we may also write $df(x)/dx$ and df/dx. The shortest notation is $f'(x)$ or y'.

It was the great Gottfried Wilhelm Leibniz who first used this elegant notation:

$$f'(x) \equiv \frac{dy}{dx} \equiv \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}. \tag{2.4}$$

Note that the limit in Eq. (2.4) is well-defined only if the curve $y = f(x)$ is continuous. Functions can only be differentiated where they are continuous.

The material world is not continuous. At small distances the discontinuity of the atomic world appears. Such discontinuities are neglected when we apply the differential calculus to the description of the world, which we do in the general theory of relativity. Then we idealize the world as a continuum. This is, however, an eminently adequate idealization for the purpose of describing most macroscopic phenomena.

The derivative, as defined in Eq. (2.4), is simply an expression of the slope of a curve. Since this is identical to the direction of the tangent at any point of a curve, the derivative is an expression of the slope of the tangent line at any point of the curve.

We shall now introduce a new concept called the differential.

In Fig. 2.5 we have drawn the graph of a function $y = f(x)$ and a tangent line of the curve in the neighbourhood of a point P with x-coordinate x_0. The tangent, which is called the linearization $L(x)$ of f, has constant slope quotient, $k = \Delta L / \Delta x$. Then $k = \Delta L / \Delta x = f'(x_0)$, since the slope of the tangent is equal to the slope of the curve at x_0. It follows that the increment of the linearization of f as x increases by Δx is $\Delta L = f'(x_0)\Delta x$. The quantity ΔL is called the differential of f and will be denoted by Df. Hence
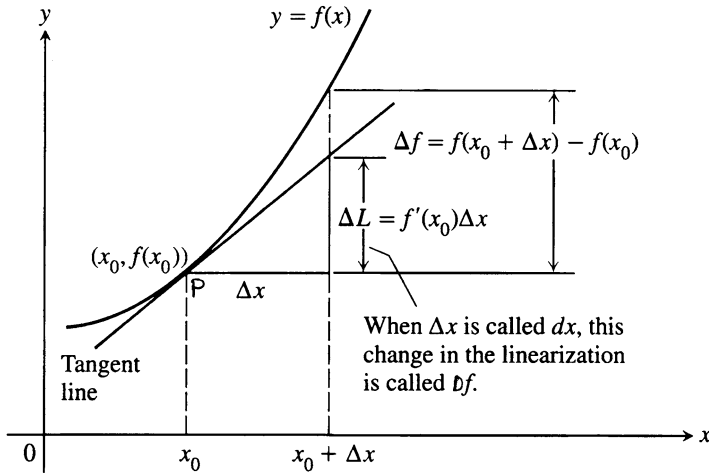
$$Df = f'(x_0)\Delta x. \tag{2.5}$$

Fig. 2.5 The tangent line at P

The increment Δf of f as x increases by Δx is $\Delta f = f(x_0 + \Delta x) - f(x_0)$. Later (see Eq. 2.93) we shall show that if Δx is small, the difference between the differential Df and the increment Δf of f is given by a sum of terms proportional to increasing powers of Δx, starting by $(\Delta x)^2$. This means that when we calculate to first order in Δx, this difference may be neglected.

2.2   Calculation of slopes of tangent lines

Let us as a first example consider the function

$$y \equiv f(x) = x^2. \tag{2.6}$$

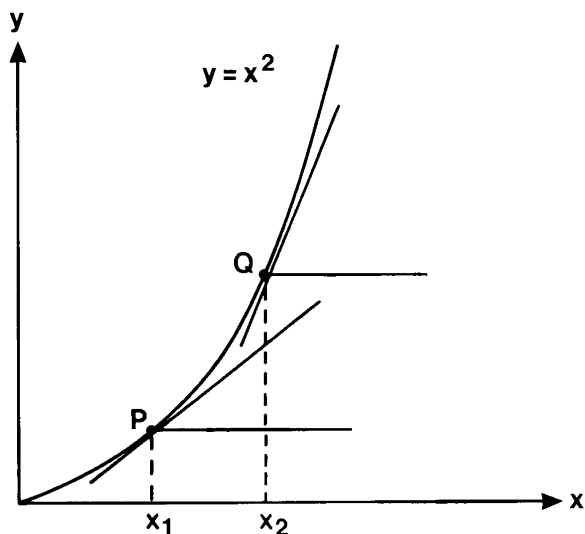Writing x_1 for $x + \Delta x$ and y_1 for $y + \Delta y$ the expression for the slope quotient takes the form

$$\frac{\Delta y}{\Delta x} = \frac{y_1 - y}{x_1 - x}. \tag{2.7}$$

According to Eq. (2.6), $y = x^2$, and $y_1 = x_1^2$. Hence we can do a first calculation, using the rule $a^2 - b^2 = (a - b)(a + b)$

$$\frac{y_1 - y}{x_1 - x} = \frac{x_1^2 - x^2}{x_1 - x} = \frac{(x_1 - x)(x_1 + x)}{x_1 - x} = x_1 + x. \tag{2.8}$$

Fig. 2.6 The slope of the
curve $y = x^2$

When x_1 approaches x, the quantity $x_1 + x$ approaches $2x$. In symbols (with $y = x^2$)

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{x_1 \to x} \frac{y_1 - y}{x_1 - x} = \lim_{x_1 \to x} (x_1 + x) = 2x. \tag{2.9}$$

From Eq. (2.9) we conclude

$$\text{if} \quad y \equiv f(x) = x^2 \quad \text{then} \quad \frac{dy}{dx} = 2x. \tag{2.10}$$

The derivative of $y = x^2$ is simply $2x$. Geometrically this means that at any point (x_1, y_1) the slope of the curve increases proportionally to x_1, as is illustrated in Fig. 2.6. Note that the derivative of a function is itself a function. When x changes, the derivative changes accordingly.

Our next example involves a little more calculation with fractions, but will be needed below. We shall find the derivative of the function $f(x) = 1/x$.

$$\left(\frac{1}{x}\right)' = \lim_{\Delta x \to 0} \frac{\frac{1}{x+\Delta x} - \frac{1}{x}}{\Delta x} = \lim_{\Delta x \to 0} \frac{\frac{x-(x+\Delta x)}{(x+\Delta x)x}}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{\frac{-\Delta x}{(x+\Delta x)x}}{\Delta x}. \tag{2.11}$$

Dividing by Δx in the upper numerator and the denominator, gives

$$\left(\frac{1}{x}\right)' = \lim_{\Delta x \to 0} \frac{-1}{(x + \Delta x)x} = -\frac{1}{x^2}. \tag{2.12}$$

As a third example let us take the derivative of the derivative of $f(x) = x^2$. This derivative is called a derivative of the second order with respect to x^2, or the second derivative of $f(x)$, and is denoted by $d^2 f/dx^2$. This time we get

$$\frac{d^2 f}{dx^2} = \lim_{\Delta x \to 0} \frac{2\left[(x + \Delta x) - x\right]}{\Delta x} = \frac{2\Delta x}{\Delta x} = 2. \tag{2.13}$$

The geometrical meaning of this result will be made clear in section 2.5. Note, however, that a 'constant function', $f(x) = k$, for instance where $k = 2$, corresponds geometrically to a horizontal straight line. There is no slope; the derivative of a constant function is zero. In the section on series expansions we shall need the following consequence of this: The third derivative of x^2 is zero, and so on for the fourth, the fifth and sixth derivative and so on.

We shall often use a couple of elementary rules:

1. The derivative of the sum of two or more functions is equal to the sum of their derivatives.
2. The derivative of the product of a constant and a function is equal to the constant times the derivative of the function.

Example 2.1.   Using the results that the derivative of x^2 is $2x$, the derivative of x is 1, and the derivative of a constant is zero, we obtain: if $y = ax^2 + bx + c$, then $dy/dx = 2ax + b$.

2.3   Geometry of second derivatives

Each curve we talk about in this section is assumed to be the graph of a function $f(x)$. The function value $f(x)$ gives the height y above the x-axis of a point on the curve $y = f(x)$. The first derivative gives the change of height per unit distance in the x-direction. Consider the function $y = 2x$. For $x = 1$ we then have $y = 2$, for $x = 2$ we have $y = 2 \times 2 = 4$, for $x = 3$ we have $y = 2 \times 3 = 6$, and so on. The change of y is the double of the change of x. The rate of change is constant. The derivative is a constant, namely 2. It determines the slope of the curve.

The second derivative gives the rate of change of the slope with distance in the x-direction. The second derivative of $y = 2x$, that is, the first derivative of 2, is zero, as it should be. But for $y = x^2$, the second derivative is not zero, but is a constant: $y' = 2x$, $y'' = 2$. The slope of the curve changes with x, and the value '2' is a measure of the change.

A straight line, i.e. a line which is not curved, has a constant slope. The more a curve curves, the faster its slope changes, and the larger is the value of the second derivative of the corresponding function $f(x)$. There is a close relation between the second derivative of a function and the curvature of its graph. This will be discussed in more detail in chapter 9.

2.4   The product rule

We shall now find a formula showing how the slope changes along the graph of a function which is the product of two other functions. Let

$$y = f(x)g(x), \quad y + \Delta y = f(x + \Delta x)g(x + \Delta x). \tag{2.14}$$

We can get an expression for Δy by subtracting y from (2.14) i.e. $\Delta y = (y + \Delta y) - y$

$$\Delta y = f(x + \Delta x)g(x + \Delta x) - f(x)g(x). \tag{2.15}$$

In order to arrive at a limit analogous to that of Eq. (2.4) we use a trick, adding and subtracting in Eq. (2.15) the same expression, that is since $-f(x)g(x + \Delta x) + f(x)g(x + \Delta x) = 0$, the right-hand side of Eq. (2.15) can be written as

$$\Delta y = f(\overset{1}{x} + \Delta x)g(x \overset{2}{+} \Delta x) - f\overset{3}{(x)}g(x \overset{4}{+} \Delta x)$$

$$+ f\overset{5}{(x)}g(x \overset{6}{+} \Delta x) - f\overset{7}{(x)}g\overset{8}{(x)}. \tag{2.16}$$

From the terms marked 1, 2, 3, and 4 we form the expression $g(x + \Delta x)[f(x + \Delta x) - f(x)]$, and from the functions 5, 6, 7, and 8 we form $f(x)[g(x + \Delta x) - g(x)]$. Dividing by Δx we get the formula

$$\frac{\Delta y}{\Delta x} = g(x + \Delta x)\frac{f(x + \Delta x) - f(x)}{\Delta x}$$

$$+ f(x)\frac{g(x + \Delta x) - g(x)}{\Delta x}. \tag{2.17}$$

From now on it is getting easier. We proceed to the limit where $\Delta x \to 0$. In the first term at the right-hand side we can then use the limiting value

$$\lim_{\Delta x \to 0} g(x + \Delta x) = g(x). \tag{2.18}$$

From the definition (2.4) we get

$$\lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{df(x)}{dx} \tag{2.19}$$

and

$$\lim_{\Delta x \to 0} \frac{g(x + \Delta x) - g(x)}{\Delta x} = \frac{dg(x)}{dx}. \tag{2.20}$$

Consequently, looking at Eqs. (2.17) and (2.18), and then at (2.17) together with (2.19) and (2.19), we find

$$\frac{d\,[f(x)g(x)]}{dx} = g(x)\frac{df(x)}{dx} + f(x)\frac{dg(x)}{dx}. \tag{2.21}$$

Perceptually clearer

$$[f(x)\,g(x)]' = g(x)f'(x) + f(x)g'(x). \tag{2.22}$$

Perceptually? Visually? What is the relevance in mathematics? Ideally no relevance, in practice quite central. A professional glances half a second at a formidable formula with 1000 signs, calmly announcing: "There is a mistake—*here*". What a formidable intellect, what a deep understanding, we are apt to think. But the expert is likely only to have activated his perceptual apparatus, nothing more. (His 'gestalt vision' I (A.N.) would say, as a philosopher.) The equation (2.21) has nearly half a hundred separate meaningful signs, Eq. (2.22) has 28. We are now offering a version of the product rule with only 12 signs. Let us pose

$$u \equiv f(x), \quad v \equiv g(x). \tag{2.23}$$

From this emerges supreme simplicity and surveyability:

$$(uv)' = vu' + uv'. \tag{2.24}$$

In words: The derivative of the product of two functions is equal to the second function multiplied by the derivative of the first function plus the first function multiplied by the derivative of the second.

Example 2.2. Find the derivative of $y = x^3$.
   Writing u for x^2 and v for x, that is uv for x^3, the use of Eqs. (2.24) and (2.10) gives us

$$(x^3)' = (x^2 \times x)' = x(x^2)' + x^2 \times x' = x \times 2x + x^2 \times 1 = 3x^2.$$

2.5   The chain rule

What is the derivative of the more complicated function $y = (x^2 + 3)^3$? Here, too, we can find a simple rule which is called 'the chain rule'. What is inside the parenthesis is itself a function of x, and we may easily perceive what is to be done by denoting it by a letter of its own, say, u. Accordingly, $y = u^3$ with $u = x^2 + 3$. The function y may then be written

$$y(u) = y[u(x)]. \tag{2.25}$$

We may think of this function in two ways; either as a function $y(u)$ of u, or as a *composite function* $y[u(x)]$ of x, i.e. a function of a function.

Let us now free ourselves of the particular example $y = (x^2 + 3)^3$ and consider two arbitrary functions $y(u)$ and $u(x)$. We shall deduce a formula for the derivative of the composite function $y[u(x)]$. Here y is called the *outer function* and u the *inner function*. According to Eq. (2.5) the differential of y is

$$Dy = y'(u) \times \Delta u, \tag{2.26}$$

where Δu is the increment of u as x increases by Δx, and the differential of the function $u(x)$ is

$$Du = u'(x) \times \Delta x. \tag{2.27}$$

As noted above the difference between Δu and Du is given by a sum of terms proportional to increasing powers of Δx, starting by $(\Delta x)^2$. We shall take the limits $\Delta u \to 0$ and $\Delta x \to 0$, so it will be sufficient to calculate to first order in Δu and Δx. Hence it is sufficient to use the approximations $Dy \approx \Delta y$ and $Du \approx \Delta u$. From Eqs. (2.26) and (2.27) we then get

$$\Delta y \approx y'(u) \times u'(x)\Delta x. \tag{2.28}$$

The increments Δy and Δx are finite quantities. Hence we may divide by Δx, and get

$$\frac{\Delta y}{\Delta x} \approx y'(u) \times u'(x). \tag{2.29}$$

Taking the limit $\Delta x \to 0$ we get

$$\lim_{dx \to 0} \frac{dy}{dx} = \lim_{dx \to 0} \frac{\Delta y}{dx} = y' = y'(u) \times u'(x). \tag{2.30}$$

This is the chain rule for differentiation of a composite function. It may be phrased as follows: The derivative of a composite function is equal to the derivative of the outer function with respect to the inner function multiplied by the derivative of the inner function.

Sometimes the notation with a quotient of two differentials is convenient when we write the derivative of a function. Using this notation the chain rule takes the form

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx} = y'(u) \times u'. \tag{2.31}$$

As a simple application of the chain rule we shall differentiate $y = (x^2 + 3)^3$. Here $y = u^3$ and $u = x^2 + 3$. Using the result of the example above with x replaced by u, we may write

$$\frac{dy}{du} = 3u^2 \tag{2.32}$$

or

$$dy = 3u^2 du. \tag{2.33}$$

If we divide each of the differentials by dx we obtain

$$y' = \frac{dy}{dx} = 3u^2 \frac{du}{dx} = 3u^2 u'. \tag{2.34}$$

Inserting the expressions for y, u, and u' we finally get

$$\left[(x^2 + 3)^3\right]' = 3\left(x^2 + 3\right)^2 2x = 6x\left(x^2 + 3\right)^2. \tag{2.35}$$

2.6 The derivative of a power function

We can use the product rule and the chain rule to find the derivative of the function $f(x) = x^p$, where p is a real number.

Equation (2.10)) and Example 2.1 suggest that $(x^n)' = nx^{n-1}$ if n is an integer number. This can easily be proved by so-called mathematical induction. The formula is clearly correct for $n = 1$, in which case $(x^1)' = x^0 \equiv 1$.

If we now assume that the formula is valid for x^{n-1}, then $(x^{n-1})' = (n-1)x^{n-2}$. By means of the product rule we get, by setting $u = x$ and $v = x^{n-1}$

$$(x^n)' = (x \times x^{n-1})' = x(n-1)x^{n-2} + 1 \times x^{n-1}$$
$$= [(n-1) + 1]x^{n-1} = nx^{n-1}. \tag{2.36}$$

We now know

1. the rule is valid for $x = 1$, and
2. if the rule is valid for x^{n-1}, then it is also valid for x^n, with n as an integer number.

These are the two criteria for the proof by mathematical induction. We have now proved the rule for the case that p is an integer number, $p = n$.

Note that by inserting $n = 3$, we get the result of the above example. Furthermore the rule is correct for $n = 0$, in which case it gives $(x^0)' = 0$, which is obviously correct since $x^0 = 1$ is a constant function. Its graph is a horizontal line with vanishing slope, i.e. the derivative of a constant is zero. From Eq. (2.12), noting that $1/x = x^{-1}$ and $1/x^2 = x^{-2}$, follows that the rule is also valid for $n = -1$, which will be used in the next section.

Is it possible to prove that the rule is valid also if p is a fraction? Consider a function $u = x^{1/n}$, where n is an integer number. This function is defined in the following way

$$\left(x^{1/n}\right)^n = x. \tag{2.37}$$

Hence

$$u^n = x. \tag{2.38}$$

Here u is a function of x. Differentiation by means of the rule (2.36) and the chain rule leads to

$$nu^{n-1} \times u' = 1. \tag{2.39}$$

Dividing by nu^{n-1} on both sides gives

$$u' = \frac{1}{nu^{n-1}}. \tag{2.40}$$

Remember also that $u^n = x$ and $u = x^{1/n}$. Then we get

$$u' = \left(x^{1/n}\right)' = \frac{1}{nu^{n-1}} = \frac{u}{nu^n} = \frac{x^{1/n}}{nx} = \frac{1}{n}x^{(1/n)-1}. \tag{2.41}$$

Thus

$$(x^p)' = px^{p-1} \tag{2.42}$$

is valid also when p is a fraction of the form $p = 1/n$, which is sufficient for our applications later on.

Let us differentiate the square root of x, which is defined by

$$\sqrt{x} \equiv x^{1/2}.$$

Using the rule (2.42), we get

$$\left(\sqrt{x}\right)' = \left(x^{1/2}\right)' = \frac{1}{2}x^{-1/2} = \frac{1}{2\,x^{1/2}} = \frac{1}{2\,\sqrt{x}}.$$

2.7 Differentiation of fractions

In chapter 44we shall need to differentiate fractions of functions. By means of the product rule, the chain rule and the rule for differentiating power functions, we shall deduce the rule for differentiating such a fraction. Let $u(x)$ and $v(x)$ be functions of x. We first consider the function

$$y[v(x)] = 1/v = v^{-1}. \tag{2.43}$$

The chain rule, in the form (2.29) with u replaced by v, gives

$$y' = y'(v) \times v'. \tag{2.44}$$

From the rule (2.42) with $p = -1$ we get

$$y'(v) = (v^{-1})' = -v^{-2} = -\frac{1}{v^2}. \tag{2.45}$$

These equations lead to

$$\left(\frac{1}{v}\right)' = -\frac{1}{v^2}v'. \tag{2.46}$$

The product rule (2.22) with $f = u$ and $g = 1/v$ gives

$$\left(\frac{u}{v}\right)' = \left(u\frac{1}{v}\right)' = \frac{1}{v}u' + u\left(\frac{1}{v}\right)'. \tag{2.47}$$

Inserting Eq. (2.46) into the last term of Eq. (2.47) we find

$$\left(\frac{u}{v}\right)' = \frac{u'}{v} + u\left(-\frac{1}{v^2} \times v'\right) = \frac{u'}{v} - \frac{uv'}{v^2}. \tag{2.48}$$

Multiplying the first term at the far right-hand side by v in the numerator and the denominator, and putting the two terms on a common denominator, we finally arrive at

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2}. \tag{2.49}$$

This is the rule for differentiating a fraction of two functions.

2.8 Functions of several variables

Geometrical pictures of functions of one variable, $y = f(x)$, are conveniently drawn as curves on an (x, y) plane. In this case the value of f is the distance (in the y direction) from the x axis to a point on the curve. But, inevitably, we have to proceed to functions of several variables. Functions with two variables, $z = f(x, y)$, can be illustrated on paper, but not so easily. They are pictured as surfaces in three-dimensional space, not curves. In this case the value of the function f is the height above the (x, y) plane of a point on the surface with coordinates (x, y, z) in the three-dimensional space (see Fig. 2.8 below).

In the above sections we have seen how the change of a function of one variable with position is described by the derivative of the function. The increase of a function $f(x, y)$ of *two* variables may be divided into two parts: The increase

$$\Delta_x f \equiv f(x + \Delta x, y) - f(x, y) \tag{2.50}$$
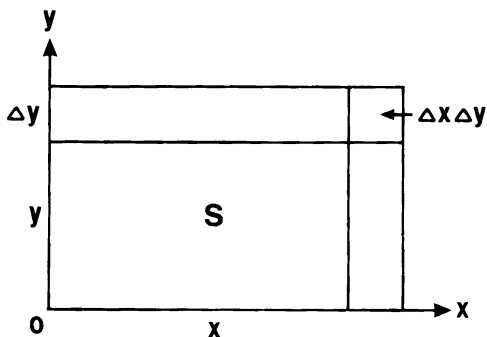
that f gets by a displacement Δx along the x-axis, and the increase

$$\Delta_y f \equiv f(x, y + \Delta y) - f(x, y) \tag{2.51}$$

it gets by a displacement Δy along the y-axis. Mathematically such variations are described by what is called 'partial derivatives' of a function of several variables.

The partial derivatives of $f(x, y)$ with respect to x and y, respectively, are defined by

$$\frac{\partial f}{\partial x} \equiv \frac{\partial f(x, y)}{\partial x} \equiv \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

$$= \lim_{\Delta x \to 0} \frac{\Delta_x f}{\Delta x} \tag{2.52}$$

and

$$\frac{\partial f}{\partial y} \equiv \frac{\partial f(x, y)}{\partial y} \equiv \lim_{\Delta y \to 0} \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y}$$

$$= \lim_{\Delta y \to 0} \frac{\Delta_y f}{\Delta y}. \tag{2.53}$$

The partial derivative of f with respect to x is calculated by differentiating f with respect to x while keeping y constant, and the partial derivative of f with respect to y is calculated by differentiating f with respect to y while keeping x constant.

We shall illustrate these new concepts by referring to an increase of the area of a rectangular garden. The Euclid-loving owner introduces a Cartesian coordinate system, as shown in Fig. 2.7. Let the function $f(x, y)$ represent the area of the garden,

$$f(x, y) = xy. \tag{2.54}$$

Let us consider two ways by which the owner may increase the area of the garden. Firstly he may extend it only in the x-direction; keeping y constant and increasing x by Δx. The corresponding increase of the area is given by Eq. (2.53),

$$\Delta_x f = (x + \Delta x) \times y - xy = xy + \Delta x \times y - xy = y\Delta x. \qquad (2.55)$$

This is just the area of the column with width Δx and height y. Secondly he may extend the garden in the y-direction; keeping x constant and increasing y by Δy. Then the increase of area is

$$\Delta_y f = x(y + \Delta y) - xy = x\Delta y. \qquad (2.56)$$

The partial derivatives of f are now found by dividing Eq. (2.50) by Δx and Eq. (2.51) by Δy, taking the limits $\Delta x \to 0$, $\Delta y \to 0$ and applying the definitions (2.52) and (2.53). The result is

$$\frac{\partial f}{\partial x} = y, \quad \frac{\partial f}{\partial y} = x \quad \text{with} \quad f(x, y) = xy. \qquad (2.57)$$

The change of height of the surface by a small displacement in arbitrary direction, with component Δx along the x-axis and component Δy along the y-axis, is

$$\Delta f = f(x + \Delta x, y + \Delta y) - f(x, y). \qquad (2.58)$$

Generalizing the definition (2.5) of the differential of a function of a single variable, we define the total differential of a function $f(x, y)$ of two variables, as

$$Df = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y. \qquad (2.59)$$

To first order in Δx and Δy there is no difference between the increment Δf and the differential Df of f. The total differential may therefore be used to calculate how the value of a function changes by small increments of the variables x and y. These increments are usually called *coordinate differentials*, and are denoted by dx and dy. With this notation the expression for a *total differential* takes the form

$$Df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy. \qquad (2.60)$$

Using Einstein's summation convention as introduced in Eq. (1.26) in Ch. 1, this may be written

$$Df = \frac{\partial f}{\partial x^i}dx^i, \qquad (2.61)$$

where $x^1 = x$ and $x^2 = y$ in the present case. (Remember that the numbers 1 and 2 are not exponents, but indices of coordinate axes.)
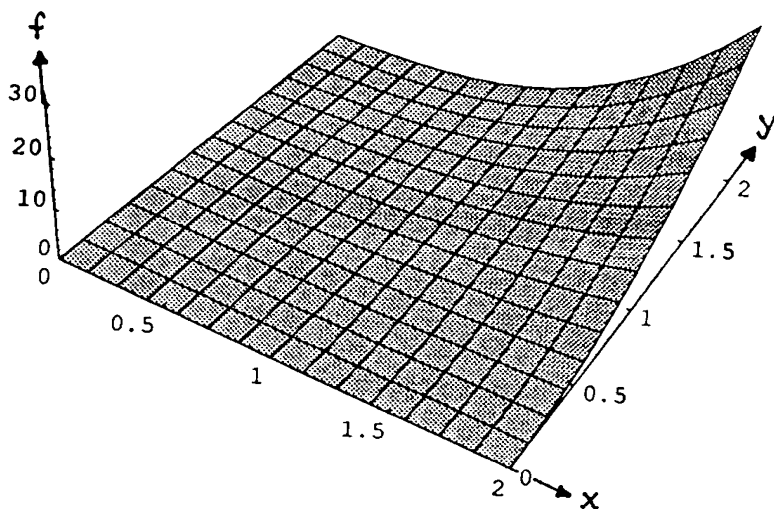
Fig. 2.8 The function $f(x, y) = x^3 y^2$

Let us illustrate the concept 'total differential' by going back to the example with the garden. The finite increment of the area of the garden is

$$
\begin{aligned}
\Delta f &= (x + \Delta x)(y + \Delta y) - xy \\
&= xy + x\Delta y + y\Delta x + \Delta x \Delta y - xy \\
&= x\Delta y + y\Delta x + \Delta x \Delta y.
\end{aligned}
\tag{2.62}
$$

Since Δx and Δy are small the product $\Delta x \Delta y$ is a 'small quantity of the second order'. In short, it is very, very small. When Δx and Δy are one to a million, $\Delta x \Delta y$ is one to a million millions. Geometrically it is the area of the small rectangle with sides Δx and Δy at the upper right-hand corner of the garden in Fig. 2.7. If we are interested in changes of the area to first order in the differentials, we can neglect products such as $\Delta x \Delta y$, which gives

$$
Df = x\,dy + y\,dx.
\tag{2.63}
$$

Further examples:

Example 2.3. We consider a hill described by the function

$$
f(x, y) = x^3 y^2.
\tag{2.64}
$$

The surface of the hill is illustrated in Fig. 2.8.

Let us first differentiate f while y is kept constant. Then we get

$$
\frac{\partial f}{\partial x} = 3x^2 y^2.
$$

Then we differentiate while x is kept constant,

$$\frac{\partial f}{\partial y} = x^3 2y = 2x^3 y.$$

Using Eq. (2.60) we get the total differential

$$Df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 3x^2 y^2 dx + 2x^3 y dy.$$

Example 2.4. Let us, as a reasonably obvious generalization, consider an example with three variables.

$$g(x, y, z) = x^2 y^3 z.$$

Then

$$\frac{\partial g}{\partial x} = 2xy^3 z, \quad \frac{\partial g}{\partial y} = 3x^2 y^2 z, \quad \frac{\partial g}{\partial z} = x^2 y^3,$$

and

$$Dg = 2xy^3 z dx + 3x^2 y^2 z dy + x^2 y^3 dz.$$

The above excessively complicated relations are found by easy, elegant manipulations of symbols. What does it mean, geometrically, or otherwise? "One cannot escape a feeling that these mathematical formulae have an independent existence and intelligence of their own, wiser than we are, ..." (Heinrich Hertz).

Einstein's field equations are 'second order partial differential equations' which contain partial derivatives of partial derivatives. Such second order partial derivatives are defined by

$$\frac{\partial^2 f}{\partial x^2} \equiv \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \frac{\partial^2 f}{\partial y \partial x} \equiv \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right),$$

$$\frac{\partial^2 f}{\partial x \partial y} \equiv \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial y^2} \equiv \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right). \qquad (2.65)$$

Just as the multiplication of 5 by 6 yields the same number as the multiplication of 6 by 5, successive differentiation with respect to x and then with respect to y, yields the same result as successive differentiation first with respect to y and then with respect to x. This is expressed mathematically by

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}, \qquad (2.66)$$

which means that different succession of differentiation does not affect the result. (The proof is of an advanced mathematical character and is not needed in our text.)

2.9   The MacLaurin and the Taylor series expansions

Most functions are more complicated and difficult to work with than power
functions. It was therefore a very useful mathematical result, when the Scottish
mathematician Colin MacLaurin (1698–1746) and the English mathematician
Brook Taylor (1685–1731) made clear that most functions can be approximated
by sums of power functions, called power series expansions. Such expansions will
be applied in chapter 9 in our discussion of curvature.

   Most anecdotes about the mathematical genius Carl Friedrich Gauss (1777–
1855) are likely to require considerable mathematical knowledge, but one is rather
innocent. Carl was a small boy when the tired teacher, hoping to get some rest,
asked his pupils to for the sum of the first hundred numbers. To his annoyance
Gauss practically at once raised his hand. He wrote from left to right 1, 2, 3, ..., 50,
then, on the next line, 100, 99, 98, ..., 51, and on the third line he added the above,
getting 50 equal sums 101, 101, 101, ..., 101. Multiplying 101 with 50 he came up
with the correct answer: 5050.

   If 100 is replaced by an arbitrary natural number n, the formula Gauss found for
the sum of the n first natural numbers, is

$$1 + 2 + 3 + \cdots + n = (1 + n)\frac{n}{2}. \tag{2.67}$$

Such a sum is called a *finite series*, meaning that it is a sum of a finite number of
terms. When the number of terms increases indefinitely, the sum of this particular
series does the same. The resulting *infinite series* is then said to *diverge*. There are,
however, series with an infinitely large number of terms which have a definite sum.
Even in the case that the sum is finite, in which case the series is said to *converge*,
such a series is called an infinite series.

   A particularly nice example of a convergent infinite series is $1 + x + x^2 + x^3 + \cdots$
with $|x| < 1$. In this sum one gets the next term multiplying the last one by x. Such a
series with infinitely many terms is called an 'infinite geometrical series'. Note that
for $|x| < 1$ each term is less than the foregoing. Let S_n denote the sum of the finite
series,

$$S_n = 1 + x + x^2 + x^3 + \cdots + x^n. \tag{2.68}$$

Multiplying the left-hand side, and each term on the right-hand side, by x,

$$x S_n = x + x^2 + x^3 + \cdots + x^n + x^{n+1}. \tag{2.69}$$

Subtracting each side of (2.68) from each side of (2.69) leads to

$$S_n - x S_n = 1 - x^{n+1} \tag{2.70}$$

or

$$(1 - x)S_n = 1 - x^{n+1}. \tag{2.71}$$

Hence

$$S_n = \frac{1 - x^{n+1}}{1 - x}. \tag{2.72}$$

If $|x| < 1$, each multiplication of x by itself gives a smaller number. Then $\lim_{n \to \infty} x^{n+1} = 0$, and the series with infinitely many terms has a finite sum equal to

$$S = \frac{1}{1 - x}. \tag{2.73}$$

The final result may be written as follows

$$1 + x + x^2 + x^3 + \cdots = \frac{1}{1 - x}, \quad |x| < 1. \tag{2.74}$$

Letting $x = 1/k$ we have an example of a convergent infinite series

$$1 + \frac{1}{k} + \frac{1}{k^2} + \cdots = \frac{1}{1 - 1/k}, \quad k > 1. \tag{2.75}$$

Here k is a real number. The left-hand side of Eq. (2.75) is meant to represent an infinite number of terms with values indicated by the first three ones that have been written down. The right-hand expression is the sum of all these terms. If $k = 2$, for example, the equation turns into

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \frac{1}{64} + \cdots$$
$$= \frac{1}{1 - 1/2} = \frac{1}{1/2} = 2. \tag{2.76}$$

Following MacLaurin and Taylor we now write a function $f(x)$ as a sum of power functions, i.e. as a *power series expansion*

$$f(x) = a_0 + a_1 x + a_2 x^2 + \cdots, \quad |x| < 1. \tag{2.77}$$

Here $a_0, a_1, a_2, \ldots$ are numbers which depend upon the function $f(x)$.

We shall now show how these numbers are determined by the values of the function and its derivatives at the point $x = 0$. Looking at Eq. (2.77) we see that if we insert the value $x = 0$ all terms on the right-hand side are equal to zero except the first one. This leads to

$$a_0 = f(0). \tag{2.78}$$

We now differentiate both sides of Eq. (2.77), using Eq. (2.42),

$$f'(x) = 1a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 + \cdots. \tag{2.79}$$

Putting $x = 0$ we get

$$a_1 = f'(0). \tag{2.80}$$

The rest of the coefficients a_2, a_3, ... can be determined in a similar way. Let us find the second and third derivatives of $f(x)$

$$f''(x) = 1 \times 2a_2 + 1 \times 2 \times 3a_3 x + 1 \times 2 \times 3 \times 4a_4 x^2 + \cdots . \tag{2.81}$$

$$f'''(x) = 1 \times 2 \times 3a_3 + 1 \times 2 \times 3 \times 4a_4 x + \cdots . \tag{2.82}$$

Putting $x = 0$ in these equations give

$$a_2 = \frac{f''(0)}{1 \times 2}, \quad a_3 = \frac{f'''(0)}{1 \times 2 \times 3}. \tag{2.83}$$

Inserting these results in Eq. (2.77) we obtain a nice formula

$$f(x) = f(0) + \frac{f'(0)}{1}x + \frac{f''(0)}{1 \times 2}x^2 + \frac{f'''(0)}{1 \times 2 \times 3}x^3 + \cdots . \tag{2.84}$$

If we continue the differentiation we can get any of the coefficients a_i expressed through derivatives of $f(x)$ at $x = 0$. The resultant infinite series is called the *MacLaurin series*.

We shall consider, as an illustration, an example that shall be employed in chapter 9; the MacLaurin series of the function $f(x) = 1/\sqrt{1 - x^2} = (1 - x^2)^{-1/2}$. In order to differentiate this function we write it as a composite function, $f[g(x)]$, where $f(g) = g^{-1/2}$, and $g(x) = 1 - x^2$. From the rule (2.42) with $p = -1/2$, and the chain rule for differentiating composite functions, we deduce

$$f'(x) = f'(g) \times g'(x) = -\left(\frac{1}{2}\right) g^{-3/2}(-2x)$$

$$= x \left(1 - x^2\right)^{-3/2}.$$

Differentiating once more, first using the product rule, and then the rule (2.42) with $p = -3/2$, we get

$$f''(x) = \left(1 - x^2\right)^{-3/2} + x \left(-\frac{3}{2}\right) \left(1 - x^2\right)^{-5/2} (-2x)$$

$$= \left(1 - x^2\right)^{-3/2} + 3x^2 \left(1 - x^2\right)^{-5/2}.$$

From these expressions follow

$$f(0) = 1, \quad f'(0) = 0, \quad f''(0) = 1.$$

Substituting this into Eq. (2.84) gives

$$\frac{1}{\sqrt{1-x^2}} = 1 + \frac{1}{2}x^2 + \cdots . \tag{2.85}$$

The third term of the series, which we have not written down, is proportional to x^4, the next one to x^6, and so on. For small x, say x smaller that 0.01, $x^4 < (0.01)^4 = 0.00000001$, showing that these higher order terms are then so small as to be negligible. For such small values of x we can use the approximation

$$\frac{1}{\sqrt{1-x^2}} \approx 1 + \frac{1}{2}x^2. \tag{2.86}$$

This approximate expression is useful because the right-hand side is easier to handle mathematically.

A second example is the MacLaurin series of the function $f(x) = \sqrt{1-x} = (1-x)^{1/2}$. Differentiation gives $f'(x) = (1/2)(1-x)^{-1/2}(-x') = (-1/2)(1-x)^{-1/2}$. Thus $f(0) = 1$ and $f'(0) = -1/2$, which gives

$$\sqrt{1-x} = 1 - \frac{1}{2}x + \cdots . \tag{2.87}$$

For small values of x we can apply the approximation

$$\sqrt{1-x^2} \approx 1 - \frac{1}{2}x^2. \tag{2.88}$$

The MacLaurin series is a series expansion of a function at the point $x = 0$, that converges for $|x| < 1$. This means that the value of the function at an arbitrary point inside the interval $-1 < x < 1$, can be found by adding terms with the values of the function and its derivatives calculated at the point $x = 0$. If, on the other hand, x is near a point outside this interval, for example $x = 2$, the MacLaurin series is of no use. Then one needs a generalization of the MacLaurin series. One needs an expansion about an arbitrary point, say $x = a$. This is called the *Taylor series*, and will be used later in our road to Einstein's field equations.

In order to arrive at the Taylor series, we introduce a function $F(x)$ defined by

$$F(x) = f(x_0 + x) \tag{2.89}$$

where x_0 is a fixed value of x. Differentiation gives

$$F'(x) = f'(x_0 + x), \quad F''(x) = f''(x_0 + x),$$
$$F'''(x) = f'''(x_0 + x), \quad \cdots . \tag{2.90}$$

Inserting $x = 0$ leads to

$$F(0) = f(x_0), \quad F'(0) = f'(x_0), \quad F''(0) = f''(x_0),$$
$$F'''(0) = f'''(x_0), \quad \cdots . \tag{2.91}$$

The MacLaurin series for the function $F(x)$ is

$$F(x) = F(0) + \frac{F'(0)}{1}x + \frac{F''(0)}{1 \times 2}x^2 + \frac{F'''(0)}{1 \times 2 \times 3}x^3 + \cdots .$$

Substituting from Eq. (2.90), we get

$$f(x_0 + x) = f(x_0) + \frac{f'(x_0)}{1}x + \frac{f''(x_0)}{1 \times 2}x^2 + \frac{f'''(x_0)}{1 \times 2 \times 3}x^3 + \cdots .$$

Adding infinitely many terms we get the Taylor series.

Substituting a coordinate increment Δx for x, the corresponding increment $\Delta f = f(x_0 + \Delta x) - f(x_0)$ of the function f is

$$\Delta f = f'(x_0)\Delta x + \frac{1}{2}f''(x_0)(\Delta x)^2 + \frac{1}{6}f'''(x_0)(\Delta x)^3 + \cdots . \tag{2.92}$$

According to Eq. (2.5) the first term at the right-hand side is just the differential Df of f. Hence

$$\Delta f - Df = \frac{1}{2}f''(x_0)(\Delta x)^2 + \frac{1}{6}f'''(x_0)(\Delta x)^3 + \cdots \tag{2.93}$$

where the terms we have not written are of higher order in Δx.

Einstein's Theory
A Rigorous Introduction for the Mathematically Untrained
Gron, O.; Naess, A.
2011, XVII, 341 p., Hardcover
ISBN: 978-1-4614-0705-8