9

Integration by steepest descents

9.1. Introduction

In some problems in radio propagation and in other branches of physics it is necessary to evaluate integrals of the form

$$I = \int_{C} g(t) \exp\{f(t)\} \mathrm{d}t \tag{9.1}$$

where C is some path in the complex plane of the variable t. This can sometimes be done by the method of steepest descents. In its simplest form it is an approximate method but it is possible to increase the accuracy by using extensions and refinements of it. One of its most important uses is in the evaluation of the contour integral representations of special functions, for example the solutions (8.16), (8.17) of the Stokes equation. These are discussed in detail in § 9.7, as an illustration of the method. The account of the method given in this chapter is only an outline. For fuller mathematical details the reader should consult a standard treatise such as Jeffreys and Jeffreys (1972) or Morse and Feshbach (1953). See also Baños (1966).

9.2. Some properties of complex variables and complex functions

A complex variable t may be expressed as the sum of its real and imaginary parts thus

$$t = x + \mathrm{i}y \tag{9.2}$$

where x and y are real. A function $f(t)$ may similarly be written

$$f(t) = \phi(x, y) + \mathrm{i}\psi(x, y) \tag{9.3}$$

where ϕ and ψ are real. Then they satisfy the Cauchy relations

$$\partial\phi/\partial x = \partial\psi/\partial y, \quad \partial\phi/\partial y = -\partial\psi/\partial x. \tag{9.4}$$

Let grad ϕ be the vector whose components are $\partial\phi/\partial x$, $\partial\phi/\partial y$ in the x–y plane, that is the complex t plane. Then (9.4) shows that

$$\operatorname{grad}\phi\cdot\operatorname{grad}\psi=0. \tag{9.5}$$

Thus the vectors grad ϕ and grad ψ are at right angles, which means that, in the complex t plane, lines on which ϕ is constant are also lines on which ψ changes most rapidly, and vice versa. The function $\exp\{f(t)\}$ has constant modulus on lines of constant ϕ, which are called 'level lines' for this function. The same function has constant argument or phase on lines for which ψ is constant. On these lines ϕ is changing most rapidly, so they are called lines of 'steepest descent' (or 'steepest ascent'). Equation (9.5) shows that the level lines and lines of steepest descent are at right angles. A diagram of the complex t plane, in which level lines for various values of $|\exp f(t)|$ are drawn, is like a contour map. The contours are the level lines, and their orthogonal trajectories are the lines of steepest descent or ascent. In this chapter frequent use is made of this kind of diagram. The value of $|\exp f(t)|$ can never be negative so that there are no contours at levels below zero. On the other hand $|\exp f(t)|$ can become indefinitely large, so that the contour maps may show some 'infinitely high' mountains. This kind of 'contour' must not be confused with a 'contour of integration', which is not, in general, a level line. In this chapter the word 'contour' is used in the sense of a level line in a contour map, and the word 'path' is used for a contour or path of integration.

If the two equations in (9.4) are again differentiated with respect to x, y (or y, x) respectively and combined, the result is

$$\frac{\partial^2\phi}{\partial x^2}+\frac{\partial^2\phi}{\partial y^2}=0,\quad \frac{\partial^2\psi}{\partial x^2}+\frac{\partial^2\psi}{\partial y^2}=0. \tag{9.6}$$

Hence both ϕ and ψ behave like potential functions in two-dimensional space, and neither can have a maximum or minimum value at any point where (9.6) holds. For example, if ϕ had a maximum at some point, this would require that $\partial^2\phi/\partial x^2$ and $\partial^2\phi/\partial y^2$ are both negative, which would violate (9.6). A point where the conditions (9.6) and the Cauchy relations (9.4) do not hold is called a 'singular point' of the function $f(t)$. A line of steepest descent is one on which ϕ decreases or increases at the greatest possible rate. It is now clear that ϕ must continue to decrease or increase until the line meets a singular point, for there is no other way in which ϕ can have a maximum or minimum value. Thus the contour maps will never show anything resembling a mountain with a top at a finite height. If a level line forms a closed curve, there must be inside this curve some singularity of $f(t)$, though not necessarily of $\exp\{f(t)\}$. This could be an infinitely high mountain, where $\phi\to+\infty$, or a zero of $\exp f(t)$, where $\phi\to-\infty$.

Clearly, a line of steepest descent (or ascent) can never form a closed curve. The term 'line of steepest descents' (plural) through a saddle point has a slightly different meaning, explained in the following section. Such a line sometimes can form a closed curve but there are no examples in this book.

9.3. Saddle points

The Cauchy relations (9.4) arise from the requirement that $f(t)$ shall have a unique derivative, which is given by

$$\frac{\mathrm{d}f(t)}{\mathrm{d}t}=\frac{\partial\phi}{\partial x}+\mathrm{i}\frac{\partial\psi}{\partial x}=\frac{\partial\psi}{\partial y}-\mathrm{i}\frac{\partial\phi}{\partial y}. \tag{9.7}$$

At a point where $\mathrm{d}f/\mathrm{d}t$ is zero, both grad ϕ and grad ψ are zero so that both the modulus and the argument of $\exp\{f(t)\}$ are stationary. Such a point is called a 'saddle point'. Let $f(t)$ have a saddle point where $t=t_0$ and let

$$f(t_0)=f_0, \quad \mathrm{d}^2f/\mathrm{d}t^2=A\mathrm{e}^{\mathrm{i}\alpha} \tag{9.8}$$

at the saddle point, where A and α are real, but f_0 is in general complex. Then Taylor's theorem gives

$$f(t)=f_0+\tfrac{1}{2}(t-t_0)^2A\mathrm{e}^{\mathrm{i}\alpha}+\mathrm{O}\{(t-t_0)^3\}. \tag{9.9}$$

Let $t-t_0=s\mathrm{e}^{\mathrm{i}\theta}$ where s and θ are real. Then

$$|\exp f(t)|=|\exp f_0|\exp\{\tfrac{1}{2}As^2\cos(2\theta+\alpha)+\mathrm{O}(s^3)\} \tag{9.10}$$

which shows that at the saddle point there are four directions (two crossing lines) for which $|\exp f(t)|$ is independent of s. These are given by

$$2\theta+\alpha=\pm\tfrac{1}{2}\pi,\ \pm\tfrac{3}{2}\pi. \tag{9.11}$$

Hence, through every saddle point there are two level lines crossing at right angles. Further

$$\arg\{\exp f(t)\}=\arg f_0+\tfrac{1}{2}As^2\sin(2\theta+\alpha)+\mathrm{O}(s^3) \tag{9.12}$$

which shows that there are four directions (two crossing lines) for which the argument or phase of $\exp f(t)$ is independent of s. These are given by

$$2\theta+\alpha=0,\ \pm\pi,\ 2\pi \tag{9.13}$$

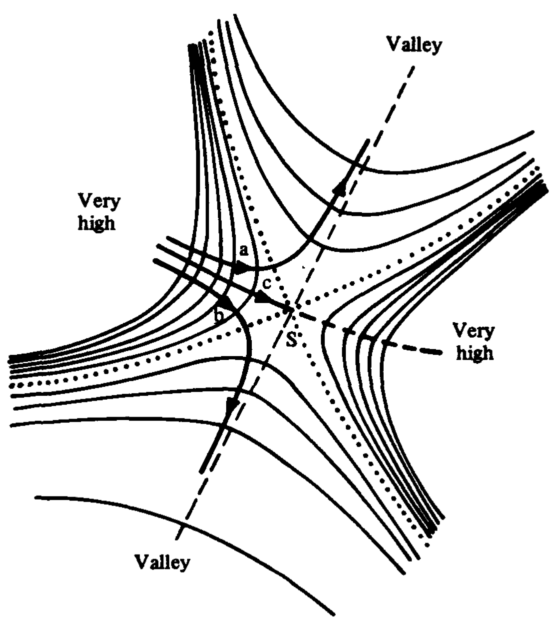
and if one of these directions is followed, in moving away from a saddle point, it becomes a line of steepest descent, or steepest ascent. Hence through every saddle point there are two such lines crossing at right angles, and their directions are at 45° to the level lines.

If any straight line through a saddle point is traversed, $|\exp f(t)|$ has a maximum or minimum value at the saddle point, according as $\cos(2\theta+\alpha)$ in (9.10) is negative or positive. If the line is tangential to a line of steepest descent or ascent, the rate of change of $|\exp f(t)|$ is greatest. In a contour map of $|\exp f(t)|$ a saddle point appears as a mountain pass, or col, shaped like a saddle.

If both $\mathrm{d}f/\mathrm{d}t$ and $\mathrm{d}^2f/\mathrm{d}t^2$ are zero at some point, then two saddle points coalesce and the point is called a 'double saddle point'. To study the behaviour of the functions, it is then necessary to include the third derivative of $f(t)$. The results are not needed in this book, but see problem 9.5.

Fig. 9.1 shows part of a contour map for $\exp\{f(t)\}$ near a saddle point. The line a is a line of steepest descent which passes very close to the saddle point but not through it. It must turn quite quickly through a right angle away from the saddle point, as shown. The line b is also a line of steepest descent which passes slightly to the other side of the saddle point; it also turns sharply through a right angle but in the opposite direction. The line c is a line of steepest descent which reaches the saddle point. To continue the descent it is necessary to turn through a right angle to either the right or left. If, instead, the line is continued straight through the saddle point, the value of $|\exp f(t)|$ is a minimum at the saddle point, and increases again when it is passed. Thus a true line of steepest descent turns sharply through a right angle when it meets a saddle point. The term 'line of steepest descents through a saddle point', however, is used in a slightly different sense to mean the line for which $|\exp f(t)|$ decreases on both sides of the saddle point at the greatest possible rate (shown as a broken line in fig. 9.1). The word 'descents' is now in the plural, because there are two descents, one

Fig. 9.1. Contour map of the region near the saddle point S. The thin continuous lines are contours, that is level lines. The thick lines are lines of steepest descent and the arrows point downhill. The dotted lines are the pair of level lines through the saddle point. The broken line is the line of steepest descents through the saddle point. Notice how the contours crowd together at the sides of the valleys.



on each side. This line passes smoothly through the saddle point and does not turn through a right angle. Crossing it at right angles at the saddle point is another line on which $|\exp f(t)|$ increases on both sides at the greatest possible rate. This is sometimes called a 'line of steepest ascents through the saddle point'.

The symbol customarily used to mark a saddle point on a contour map can be seen at S_1 and S_2 in figs. 9.5–9.10. The two reversed arcs enclose the line of steepest descents.

9.4. Error integrals and Fresnel integrals

The integral

$$I_0 = \int_{-\infty}^{\infty} \exp(-t^2)\mathrm{d}t = \pi^{\frac{1}{2}} \tag{9.14}$$

may be thought of as a contour integral in the complex t plane, in which the path is the real axis. The integrand is greatest near $t = 0$ and decreases rapidly as $|t|$ increases. When $|t|$ is large the integrand is very small, so that if the infinite limits are replaced by finite values $\pm L$, the error is small when L is large. The value of the integral when modified in this way is

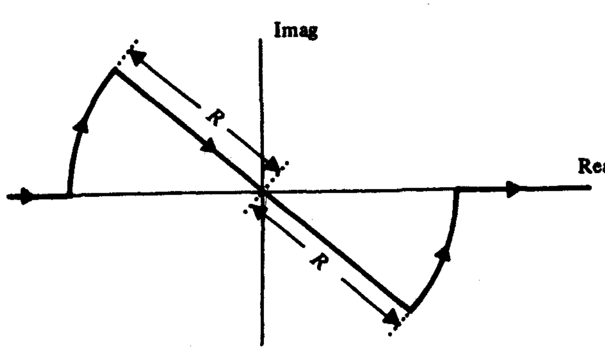
$$\int_{-L}^{L} \exp(-t^2)\mathrm{d}t = 2\int_{0}^{L} \exp(-t^2)\mathrm{d}t = \pi^{\frac{1}{2}}\,\mathrm{erf}(L) \tag{9.15}$$

where erf denotes the 'error function'. The fractional error introduced by the modification is therefore

$$1-\mathrm{erf}(L) = \mathrm{erfc}(L) = 2\pi^{-\frac{1}{2}}\int_{L}^{\infty} \exp(-t^2)\,\mathrm{d}t. \tag{9.16}$$

The error function and its complement erfc are tabulated (for example by Abramowitz and Stegun, 1965, pp. 310 ff.). The tables show that (9.16) is about 1/100 when $L = 1.82$ and about 1/10 when $L = 1.16$. This idea that the main contribution

Fig. 9.2. The complex t plane. Path of integration for the integral I_0.



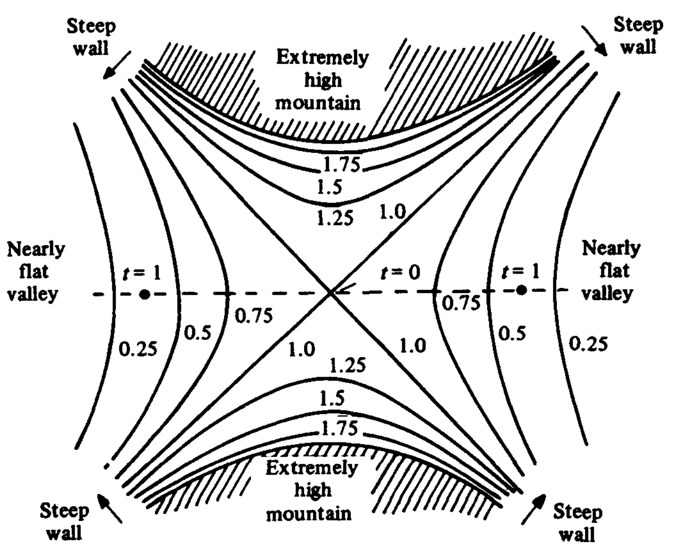
to I_0 comes from small values of $|t|$ less than about 2 is of great importance in the method of steepest descents.

The integrand of I_0 has no singularities (except at ∞) so the path can be distorted, without affecting the value of the integral, provided that it begins and ends as required by the limits of the integral. For example it can be formed by a straight line through the origin with $\arg t = -\frac{1}{4}\pi$, as shown in fig. 9.2 provided that the ends of this line are joined to the real axis. This can be done by arcs of circles of large radius R. We have already seen that the remote parts of the real axis contribute very little to the integral, and it can be proved that their contribution tends to zero as R tends to infinity. Consider now the contribution to I_0 from the 45° circular arc from $\arg t = -\frac{1}{4}\pi$ to 0. Let $t = R\exp(\mathrm{i}\theta)$. Then the contribution is

$$\mathrm{i}R\int_{-\frac{1}{4}\pi}^{0} \exp\{-R^2(\cos 2\theta + \mathrm{i}\sin 2\theta) + \mathrm{i}\theta\}\,\mathrm{d}\theta. \tag{9.17}$$

Now for $-\frac{1}{4}\pi \leqslant \theta \leqslant 0$ it is easy to show that $\cos 2\theta \geqslant 1 + 4\theta/\pi$ so that the modulus of the integrand of (9.17) is less than or equal to $\exp\{-R^2(1 + 4\theta/\pi)\}$. This expression can be integrated with respect to θ and the result shows that the modulus of (9.17) is less than $\{1 - \exp(-R^2)\}\pi/4R$. But this tends to zero as R tends to infinity so that the contribution (9.17) also tends to zero. The same applies to the other 45° arc in fig. 9.2. Thus the line at $\theta = -\frac{1}{4}\pi$ in fig. 9.2 could be supposed to extend to infinity. It need not return to the real axis. The limit $+\infty$ in I_0 (9.14) can be supposed to mean that $|t| \to \infty$ with $-\frac{1}{4}\pi \leqslant \arg t \leqslant \frac{1}{4}\pi$, and similarly for the limit $-\infty$.

Fig. 9.3. The complex t plane showing a contour map for the function $\exp(-t^2)$. The numbers by the curves are the values of $\exp(-t^2)$. The arrows point downhill.

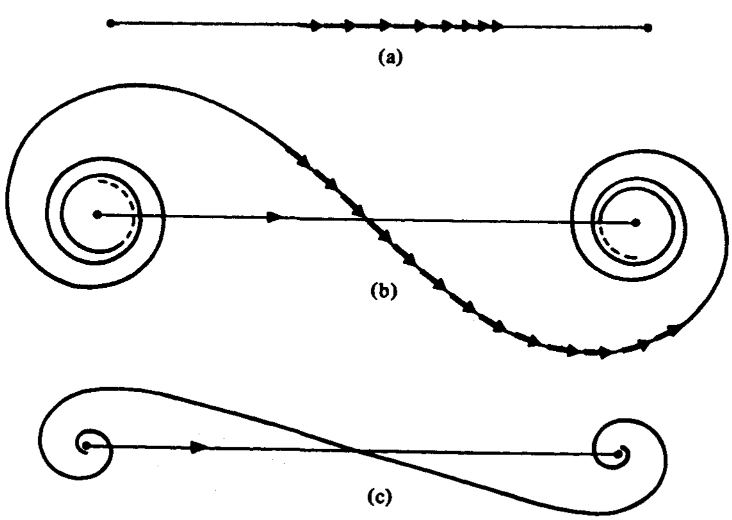


The contours or level lines for the integrand $\exp(-t^2)$ of I_0 are shown in fig. 9.3. It is clear that where $|t|$ is large there are wide valleys in which the level is almost zero, where $-\frac{1}{4}\pi < \arg t < \frac{1}{4}\pi$ and $\frac{3}{4}\pi < \arg t < \frac{5}{4}\pi$. The sides of these valleys are near the lines $\arg t = \frac{1}{4}\pi, \frac{3}{4}\pi, \frac{5}{4}\pi, \frac{7}{4}\pi$, and here the gradient is extremely steep and gets steeper as $|t|$ increases. For large $|t|$ the sides of the valleys are almost like vertical walls of extremely great height. The two ends of the path of integration for I_0 can be anywhere at infinity in these two valleys, but the walls form a boundary, beyond which the ends cannot go. In the sectors $\frac{1}{4}\pi < \arg t < \frac{3}{4}\pi$ and $\frac{5}{4}\pi < \arg t < \frac{7}{4}\pi$ there are extremely high mountains.

A complex integral like I_0 can be thought of as the sum of many complex numbers $\exp(-t^2)\mathrm{d}t$ each of which can be represented as a vector in the Argand diagram. This gives a diagram sometimes called an 'amplitude–phase' diagram or vector diagram, of the kind used in the study of optical diffraction (for examples see figs. 8.12, 8.13, 9.4). The value of the integral is then the 'resultant' or closing vector of this diagram. The form of the diagram depends on the choice of path. The resultant must be unaffected by the choice of path because the integrand is analytic.

Suppose now that the path for I_0 is the real t axis. This is the line of steepest descents through the saddle point, $t = 0$, of the integrand $\exp(-t^2)$. The vectors $\exp(-t^2)\mathrm{d}t$ now all have zero argument or phase, but for equal elements $|\mathrm{d}t|$ of the path their magnitudes decrease rapidly as $|t|$ increases. The vector diagram is therefore a straight line, fig. 9.4(a), whose length is $\pi^{\frac{1}{2}}$. Next suppose that the path is

Fig. 9.4. Amplitude–phase diagrams for various different choices of integration path in the complex t plane for the integral I_0, (9.14). (a) The real axis or line of steepest descents. Parallel vectors not of equal length. (b) A level line through the saddle point. Vectors of equal length but variable direction, giving Cornu spiral. (c) Intermediate case. Vectors vary in both length and direction.



the line $\arg t = -\frac{1}{4}\pi$. Let

$$t = s\exp(-\tfrac{1}{4}\mathrm{i}\pi) \tag{9.18}$$

so that $\mathrm{d}t = \mathrm{d}s\exp(-\frac{1}{4}\mathrm{i}\pi)$, where s is real. The path is now a level line through the saddle point at $t = 0$. The integrand is $\exp(\mathrm{i}s^2 - \frac{1}{4}\mathrm{i}\pi)\mathrm{d}s$. For equal elements $\mathrm{d}s$ the vectors in the vector diagram now all have the same length, $\mathrm{d}s$, but their phase is $\varphi = s^2 - \frac{1}{4}\pi$. The amplitude–phase diagram is a Cornu spiral, as shown in fig. 9.4(b). For small s the phase φ is near $-\frac{1}{4}\pi$ and changes only slowly with s, so that the vectors near the middle of the diagram have nearly constant direction. The direction of the change reverses where $s = 0$, for then $\mathrm{d}\varphi/\mathrm{d}s = 0$, and this is called a point of stationary phase. For large $|s|$, however, the phase changes very rapidly and then the arms of the spiral are tightly wrapped. They must converge to limit points in such a way that the resultant has length $\pi^{\frac{1}{2}}$ and argument zero, as before.

Finally suppose that the path is the line $\arg t = -\gamma$ where $\gamma \neq 0$ or $\frac{1}{4}\pi$. This still goes through the saddle point but is neither a level line nor a line of steepest descent. Clearly now for constant $|\mathrm{d}t|$ the vectors $\exp(-t^2)\mathrm{d}t$ change both in magnitude and in phase as t varies. The vector diagram is a double spiral which converges to the limit points more rapidly than the Cornu spiral. This is illustrated in fig. 9.4(c) which shows the spiral for $\gamma \approx \frac{1}{8}\pi$. Again the resultant must have length $\pi^{\frac{1}{2}}$ and argument zero. Other paths could be chosen which are not straight and do not pass through the saddle point. They could give very complicated vector diagrams. These paths must begin and end in the correct valleys, and then the resultant vector must in all cases be the same.

Consider again the integral I_0 when the path is the level line as given by (9.18). Then the integral is

$$\begin{aligned} I_0 = \pi^{\frac{1}{2}} &= \exp(-\tfrac{1}{4}\mathrm{i}\pi)\int_{-\infty}^{\infty}\exp(\mathrm{i}s^2)\mathrm{d}s \\ &= (2\pi)^{\frac{1}{2}}\exp(-\tfrac{1}{4}\mathrm{i}\pi)\left[\int_0^{\infty}\cos(\tfrac{1}{2}\pi u^2)\mathrm{d}u + \mathrm{i}\int_0^{\infty}\sin(\tfrac{1}{2}\pi u^2)\mathrm{d}u\right]. \end{aligned} \tag{9.19}$$

The two integrals in square brackets may be written $\mathrm{Lim}_{x\to\infty}C(x)$ and $\mathrm{Lim}_{x\to\infty}S(x)$ respectively where

$$C(x) = \int_0^x \cos(\tfrac{1}{2}\pi u^2)\mathrm{d}u \quad S(x) = \int_0^x \sin(\tfrac{1}{2}\pi u^2)\mathrm{d}u. \tag{9.20}$$

These are the Fresnel integrals. The locus of a point in the C–S plane is a Cornu spiral which converges to the limit points $C = S = \frac{1}{2}$, and $C = S = -\frac{1}{2}$, when $x \to \pm\infty$. The two functions (9.20) are tabulated (Abramowitz and Stegun, 1965). If the limits $\pm\infty$ in (9.19) are replaced by $\pm L$ (compare (9.15)) the fractional error can be calculated from these tables and it is found to be about 1/100 when $L = 45$ and about 1/10 when $L = 4.5$ (compared with 1.82 and 1.16 for (9.15)). Thus the form (9.19) with the level line path converges much more slowly than the form (9.14) with the path of steepest descents. This shows why the path of steepest descents has special importance.

9.5. Contour maps

In the contour map for $\exp(-t^2)$, fig. 9.3, the line of steepest descents through the saddle point and the level lines through the saddle point are straight. On either level line the slope of the hillside increases as distance from the saddle point increases. The valleys, which are nearly flat and at nearly zero altitude when $|t|$ is large, extend on the two opposite sides of the saddle point. They contain the line of steepest descents and they are bounded by the level lines. The contour map for $\exp(-\beta t^2)$ has similar features. If the constant β is real and positive the only change is a change of scale of the map by a factor $\beta^{\frac{1}{2}}$. If β is complex, the line of steepest descents through the saddle point $t=0$ is where $\arg t = -\frac{1}{2}\arg\beta$, and the level lines through the saddle point are where $\arg t = -\frac{1}{2}\arg\beta \pm \frac{1}{4}\pi$. Thus the contour map is obtained from that for $\exp(-t^2)$ simply by a rotation through an angle $-\frac{1}{2}\arg\beta$ and a scale change by a factor $|\beta|^{\frac{1}{2}}$.

For a more general function $f(t)$, the contour map for $\exp\{f(t)\}$ has more than one saddle point in nearly all cases of practical importance. The contour map near one saddle point S_1 that is remote from singularities of $f(t)$ and from other saddle points is similar to that for $\exp(-\beta t^2)$, as was described in § 9.3. But now, in general, the level lines and lines of steepest descents through S_1 are curved; see fig. 9.1 for example. The most important feature still is that valleys extend in two sectors from S_1. They are bounded by the level lines through S_1 and they contain the line of steepest descents.

Now consider another saddle point S_2 that is the nearest to S_1. There are several possibilities. (a) If S_2 is where $|\exp\{f(t)\}|$ is greater than at S_1, it is in one of the two mountain regions extending from S_1, and the line of steepest descents through S_1 cannot go near it. (b) If S_2 is where $|\exp\{f(t)\}|$ is less than at S_1 it must be in one of the valleys extending from S_1. Adjoining S_2 are two mountain regions and S_1 is in one of these. The other, remote from S_1, must have a contour at the same level as S_1 because all mountains are infinitely high (see § 9.2). Beyond this contour there might be other saddle points or singularities. This mountain divides the valley of S_1 into two parts, V_1 and V_2. The line of steepest descents for S_1 runs down into one part, say V_1. When using an integration path that is required to run down into V_2, we let the path first run from S_1 along the line of steepest descents into valley V_1 and thence through S_2 along its line of steepest descents into valley V_2. Thus there is a contribution to the integral from near S_2 as well as from near S_1. If a path from S_1 running directly into V_2 were used, the integrand could not decrease as fast as it does on the line of steepest descents, so that there are contributions to the integral not only from near S_1 but from more remote parts of the path. These contributions are taken care of by first using the line of steepest descents for S_1 and then adding in the contribution from near S_2. (c) If S_2 and S_1 are at the same level, this is really a special case of (b). The steepest descents contributions from S_1 and S_2 are now of comparable order of magnitude. An example of this case is given in § 9.7, fig. 9.11.

For further understanding of how to choose the integration path it is helpful to study actual examples. These can be found in §§ 9.7, 10.18–10.20.

9.6. Integration by the method of steepest descents

Consider first a simplified form of (9.1) in which $g(t) = 1$ so that

$$I = \int_C \exp\{f(t)\}\mathrm{d}t. \tag{9.21}$$

To evaluate this there are two steps. First, the path C is deformed so that it goes through one or more saddle points of the integrand into the valleys on either side. Then the contributions to the integral from near each used saddle point are evaluated and added.

In problems of physics it nearly always happens that each end of the path C is at a singularity of $f(t)$ where

$$\mathrm{Re}\{f(t)\} \to -\infty, \ |\exp\{f(t)\}| \to 0. \tag{9.22}$$

Then the path begins and ends in a valley. To deform the path we first find the valleys for all the saddle points of $f(t)$. Then from the valley where C begins we let the path run over a saddle point along the line of steepest descents into another valley. If the end point of C is not in this valley we now go along the line of steepest descents of another saddle point into a third valley. The process is repeated until we reach the valley where the end point of C is known to lie. Thus the path is distorted to follow one or a succession of lines of steepest descents through saddle points. This process is best illustrated by the specific examples given later.

Suppose that one of the saddle points on the deformed path is at t_0. On the line of steepest descents through it, $\mathrm{Im}\{f(t)\}$ is constant, so the equation of this line is

$$\mathrm{Im}\{f(t)\} = \mathrm{Im}\{f(t_0)\}. \tag{9.23}$$

This is also the equation of the line of steepest ascents through t_0, and it may have other branches which do not pass through t_0. The branch of (9.23) that is to be used for the path must be found by inspection (see § 9.7, for some examples), but as already explained it is not necessary to know its exact position.

When the path of integration is moved, certain precautions must be taken. The following remarks apply also for the more general form (9.1) with a factor $g(t)$ in the integrand.

If, in moving the path of integration, a pole of the integrand is crossed, it must be allowed for. If it is remote from the contributing saddle points, this can be done by appropriately adding or subtracting $2\pi i$ times its residue. If it is near a saddle point, a process for removing the pole must be used before applying the method of steepest descents (Baños, 1966). If a branch point is crossed, a suitable branch cut must be inserted, and the contribution from the integral around this cut must be included. Both these cases occur in radio propagation problems, the pole in the theory of surface waves (Sommerfeld, 1909; Baños, 1966), and the branch cut in the theory of guided waves and lateral waves (Ott, 1942; Felsen, 1967) but they are beyond the scope of this book.

The contribution to the integral (9.21) from near t_0 is now found as follows. The exponent $f(t)$ is expanded in a Taylor series about t_0, so that the integrand becomes

$$\exp\{f(t)\} = \exp(f_0)\exp\left\{\frac{1}{2!}(t-t_0)^2 f_0'' + \frac{1}{3!}(t-t_0)^3 f_0''' + \cdots\right\} \tag{9.24}$$

where f_0, f_0'', f_0''' are the values of f, $\mathrm{d}^2f/\mathrm{d}t^2$, $\mathrm{d}^3f/\mathrm{d}t^3$ at the saddle point. On the line of steepest descents through t_0 the last exponent (the series) in (9.24) is real because of (9.23). It is zero at t_0 and negative elsewhere. Hence let

$$\frac{1}{2!}(t-t_0)^2 f_0'' + \frac{1}{3!}(t-t_0)^3 f_0''' + \cdots = -s^2 \tag{9.25}$$

where s is real. The sign of s is to be chosen so that s is negative when the point t moves along the path towards the saddle point, and positive after t has passed the saddle point and is moving away from it. At the two ends of the line of steepest descents $\mathrm{Re}\{f(t)\} = -\infty$ so that s runs from $-\infty$ to $+\infty$. Now if $|t-t_0|$ is small enough the second and later terms of (9.25) are small compared with the first. Thus

$$\mathrm{d}t \approx \{f_0''/2!\}^{-\frac{1}{2}}\mathrm{i}\,\mathrm{d}s. \tag{9.26}$$

Let

$$f_0'' = A\mathrm{e}^{\mathrm{i}\alpha} \quad (0 \leqslant \alpha < 2\pi). \tag{9.27}$$

Then (9.26) gives

$$\mathrm{d}t = \pm(2/A)^{\frac{1}{2}}\exp\{\mathrm{i}\tfrac{1}{2}(\pi-\alpha)\}\,\mathrm{d}s = P\,\mathrm{d}s. \tag{9.28}$$

The sign depends on the angle Θ between the path and the real t axis at the saddle point. It is positive if $-\frac{1}{2}\pi < \Theta \leqslant \frac{1}{2}\pi$ and negative otherwise.

The integrand of (9.21) contains the factor $\exp(-s^2)$ which is appreciable only near $s=0$, that is when $|t-t_0|$ is small. If we use the approximation (9.28), the contribution to the integral from the neighbourhood of t_0 is

$$\exp(f_0)P\int_{-\infty}^{\infty}\exp(-s^2)\,\mathrm{d}s = \pm(2\pi/A)^{\frac{1}{2}}\exp(f_0)\exp\{\tfrac{1}{2}\mathrm{i}(\pi-\alpha)\}. \tag{9.29}$$

The sign depends on Θ and must be decided by inspection of the path (see the examples in § 9.7).

Formula (9.29) is the main result of the method of steepest descents. It is approximate because of (9.26) but the error is often very small. More accurate formulae are derived later, § 9.9, and these provide an estimate of the error. It can be seen from (9.25), however, that since the term $\frac{1}{6}f_0'''(t-t_0)^3$ and higher terms have been neglected, they ought to be small compared with $\frac{1}{2}(t-t_0)^2 f_0''$ for values of s up to about 2; see section 9.4 just after (9.16). We shall expect the method to fail if f_0'' is too small. But this case can be dealt with by using Airy integral functions, which is perhaps the most important use of these functions; §§9.9 (end) and 10.20.

Consider now the more general integral (9.1) with the factor $g(t)$ restored. It can be written

$$\int_C \exp\{f(t) + \ln g(t)\}\,\mathrm{d}t \tag{9.30}$$

and the foregoing method could in principle be applied to it. Complications arise however, if $g(t)$ has a pole or a zero near a saddle point of $\exp\{f(t)\}$. It commonly happens that $g(t)$ is a 'slowly varying' function of t near a saddle point of $\exp\{f(t)\}$. Then in the region of the t plane where the integrand of (9.1) is appreciable $g(t)$ may be taken as a constant, $g(t_0)$. The contribution to (9.1) from near the saddle point t_0 of $\exp\{f(t)\}$ is then

$$\pm(2\pi/A)^{\frac{1}{2}} g(t_0)\exp(f_0)\exp\{\tfrac{1}{2}\mathrm{i}(\pi-\alpha)\}. \tag{9.31}$$

When $g(t)$ cannot be treated as constant, it is expanded in powers of s and included as a factor in the integrand of (9.29). If terms after the first on the left of (9.25) can be neglected, s is proportional to $t-t_0$, so that $g(t)$ is simply expanded in powers of $t-t_0$:

$$g(t) = g(t_0) + (t-t_0)g_0' + \frac{1}{2!}(t-t_0)^2 g_0'' + \ldots \tag{9.32}$$

where $g_0', g_0'' \cdots$ are $\mathrm{d}g/\mathrm{d}t$, $\mathrm{d}^2 g/\mathrm{d}t^2 \cdots$ at $t=t_0$. On integration, the odd powers of s or $t-t_0$ give zero. The third term of (9.32) gives a contribution which must be added to (9.31), and when (9.27) is used, this is

$$\begin{aligned} &-(g_0''/f_0'')\exp(f_0)P\int_{-\infty}^{\infty}\exp(-s^2)s^2\,\mathrm{d}s \\ &= \mp\tfrac{1}{2}g_0''(2\pi)^{\frac{1}{2}}A^{-\frac{3}{2}}\exp(f_0)\exp\{\tfrac{1}{2}\mathrm{i}(\pi-3\alpha)\}. \end{aligned} \tag{9.33}$$

This term may be important if $g(t)$ has a zero at or near t_0, so that $g(t_0)$ in (9.31) is zero or small.

When terms after the first, on the left of (9.25), cannot be neglected, the more elaborate treatment of § 9.9 must be used.

For the methods of (9.31)–(9.33) to work there is one fairly obvious restriction on the function $g(t)$. Suppose that the saddle points and lines of steepest descents for $\exp\{f(t)\}$ have been found. On these lines this function contains a factor $\exp(-s^2)$

whose properties are used in the integral (9.29). But this would be wrong if $g(t)$ contained a factor that behaved like $\exp(Ks^2)$ where K is real and greater than unity. The path would not now be on a line of steepest descents and the extra factor would make the integral in (9.29) diverge. Thus $g(t)$ must be 'devoid of exponential behaviour' (quoted from Baños, 1966). In fact the extra factor would not matter if K were less than unity, though its presence might reduce the accuracy of the steepest descents result.

9.7. Application to solutions of the Stokes equation

The simplest possible integral that can be evaluated by steepest descents is the error integral, I_0 in (9.14). When the integrand is the exponential of any quadratic function, it can be changed to the form I_0 simply by a change of variable. There is only one saddle point, where $t=0$, and the value $\pi^{\frac{1}{2}}$ in (9.14) is exact.

The next simplest possible integral is when the integrand is the exponential of a cubic function of t. This case is therefore very useful for illustration. It is very important in physics because the solutions (8.16), (8.17) of the Stokes differential equation can be expressed in this form. We shall therefore study the integral (8.16) for $\mathrm{Ai}(\zeta)$.

The exponent in (8.16) has two saddle points where

$$t = t_0 = \pm \zeta^{\frac{1}{2}} \tag{9.34}$$

and the lines of steepest descents through them will first be found. The path of integration is then distorted to coincide with one or both of them. The configuration of these lines depends on the value of $\arg \zeta$, and it will be shown that sometimes one and sometimes both of these lines must be used. The transition from one case to the other gives the Stokes phenomenon, § 8.12.

Let $t = u + \mathrm{i}\, v$ where u and v are real and let

$$\zeta = m^2 \mathrm{e}^{\mathrm{i}\theta} \tag{9.35}$$

where m and θ are real and positive. Then the two saddle points are where

$$t = t_0 = \pm m \exp(\tfrac{1}{2}\mathrm{i}\theta). \tag{9.36}$$

The exponent $\zeta t - \frac{1}{3}t^3$ in (8.16) at either saddle point has the value

$$f_0 = \tfrac{2}{3}\zeta t_0 \tag{9.37}$$

and its second derivative is

$$f''_0 = -2t_0. \tag{9.38}$$

On the lines of steepest descent and ascent through a saddle point the imaginary part of the exponent is constant. Hence the equation of these lines is

$$\mathrm{Im}(\zeta t - \tfrac{1}{3}t^3) = \mathrm{Im}(f_0) \tag{9.39}$$

(compare 9.23) which leads to

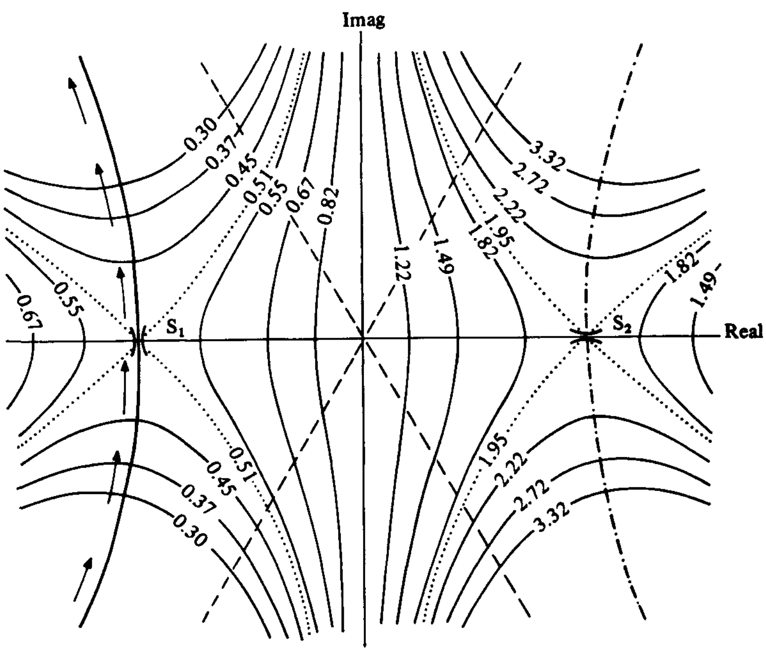
$$\tfrac{1}{3}v^3 - u^2v + vm^2\cos\theta + um^2\sin\theta = \pm\tfrac{2}{3}m^3\sin(\tfrac{3}{2}\theta). \tag{9.40}$$

Suppose first that ζ is real and positive so that $\theta = 0$. Then the saddle points lie on the real t axis at $\pm m$ and (9.40) becomes

$$v(\tfrac{1}{3}v^2 - u^2 + m^2) = 0. \tag{9.41}$$

The curves are shown in fig. 9.5 which is a contour map of the integrand. The line $v = 0$ is a line of steepest ascent through the saddle point S_1 and a line of steepest descent through the saddle point S_2. The other factor is the equation of a hyperbola whose asymptotes are given by $v = \pm u\sqrt{3}$. The left branch is a line of steepest descents through the saddle point S_1, and it begins and ends where $\arg t \approx \mp\tfrac{2}{3}$, that is within the valleys where the path C must begin and end. Hence the path of integration may be distorted to coincide with the left-hand branch of the hyperbola as indicated by arrows in fig. 9.5. The right-hand branch of the hyperbola, shown as a chain line in fig. 9.5, is a line of steepest ascents through the other saddle point S_2,

Fig. 9.5. Complex t plane. Contour map of the function $\exp(t\zeta - \tfrac{1}{3}t^3)$, when ζ is real and positive. The thin continuous lines are contours, and the numbers on them are the values of the altitude $\exp\{\mathrm{Re}(t\zeta - \tfrac{1}{3}t^3)\}$. In this example and in figs. 9.6–9.11, $|\zeta| = 1.0$. The thick continuous line is the line of steepest descents through the saddle point S_1 and the arrows show the path for the contour integral (8.16). The broken lines are at $\pm 60°$ to the real axis and are asymptotes of the lines of steepest descent and ascent.



and is not part of the path, which therefore passes through only the one saddle point S_1 at $t_0 = -m$. Equations (9.37), (9.38) then show that $f_0 = -\frac{2}{3}\zeta^{\frac{3}{2}}$ and $f_0'' = 2\zeta^{\frac{1}{2}}$, so in (9.27), $\alpha = 0$, $A = 2\zeta^{\frac{1}{2}}$, where the positive value of $\zeta^{\frac{1}{2}}$ is used.

The formula (9.29), with the factor $1/(2\pi i)$ from (8.16), may now be used to give the approximate value of the integral. The direction Θ of the contour is $\frac{1}{2}\pi$ so the sign in (9.28) is $+$, whence

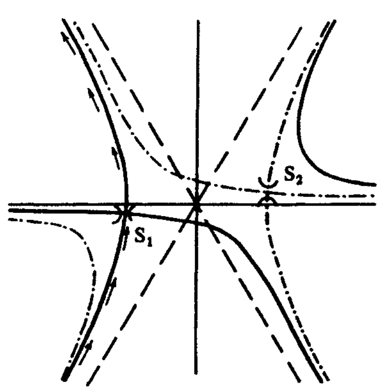
$$\mathrm{Ai}(\zeta) \sim \tfrac{1}{2}\pi^{-\frac{1}{2}}\zeta^{-\frac{1}{4}}\exp(-\tfrac{2}{3}\zeta^{\frac{3}{2}}) \tag{9.42}$$

where the real positive values of $\zeta^{-\frac{1}{4}}$ and $\zeta^{\frac{3}{2}}$ are used. This is the asymptotic form for $\mathrm{Ai}(\zeta)$ when ζ is real and positive (compare (8.52)). It is the same as the subdominant W.K.B. term (8.27) with a multiplying constant. This illustrates a most important use of the first order steepest descents formula (9.29). When applied to a contour integral representation of a solution of a differential equation, the contribution from one saddle point, if sufficiently isolated, is one of the W.K.B. solutions.

The unused saddle point S_2 in fig. 9.5 is at a much greater altitude than S_1. For other solutions of the Stokes equation, for example $\mathrm{Bi}(\zeta)$, it does give a contribution, proportional to the dominant term, (8.28).

Next suppose that $\theta(=\arg\zeta)$ is in the range $0 < \theta < \frac{1}{3}\pi$. The expression (9.40) now gives two separate equations. The plus sign gives the lines of steepest ascent and descent through the saddle point S_2 at $t_0 = me^{\frac{1}{2}i\theta}$, shown as a chain line in fig. 9.6, and the minus sign gives those through the other saddle point S_1, shown as continuous lines. Each equation is now a cubic. The curve for the minus sign has two intersecting branches which are the lines of steepest descent and ascent through the saddle point S_1 at $t_0 = -me^{\frac{1}{2}i\theta}$, and a third branch which does not pass through either saddle point. The asymptotes are the lines $v = 0$, $v = \pm u\sqrt{3}$. The line of steepest descents begins and ends in the correct sectors as before, and is used as the path C_1, which is indicated by arrows in fig. 9.6, and which therefore passes

Fig. 9.6. Similar to fig. 9.5 but for $0 < \arg\zeta < \frac{1}{3}\pi$. In this and figs. 9.7–9.11 the contours (level lines) are not shown.



through only the one saddle point at $t_0 = -me^{\frac{1}{2}i\theta}$. Equations (9.37) and (9.38) show that $f_0 = -\frac{2}{3}\zeta^{\frac{3}{2}}$ (where arg $\zeta^{\frac{3}{2}}$ is $\frac{3}{2}\theta$) and $f_0'' = 2me^{\frac{1}{2}i\theta}$, so that $\alpha = \frac{1}{2}\theta$, $A = 2m$. The direction Θ of the path at the saddle point is $\frac{1}{2}\pi - \frac{1}{4}\theta$, so the + sign is used in (9.29) which gives

$$\mathrm{Ai}(\zeta) = \tfrac{1}{2}\pi^{-\frac{1}{2}}|m^{-\frac{1}{2}}|e^{-\frac{1}{4}i\theta}\exp(-\tfrac{2}{3}\zeta^{\frac{3}{2}}). \tag{9.43}$$

This is the same as (9.42) if $\zeta^{-\frac{1}{4}}$ takes the value which goes continuously into $|\zeta^{-\frac{1}{4}}|$ as θ decreases to zero.

The curves given by (9.40) can be drawn in a similar way for other values of $\theta = \arg\zeta$. Fig. 9.7 shows them for $\theta = \frac{1}{3}\pi$, which is an anti-Stokes line, and fig. 9.8 shows them for $\frac{1}{3}\pi < \theta < \frac{2}{3}\pi$. In both these cases the path is similar to those in figs. 9.5 and 9.6, and the approximate value of the integral is given by (9.42) as before. In

Fig. 9.7. Arg $\zeta = \frac{1}{3}\pi$. Anti-Stokes line.

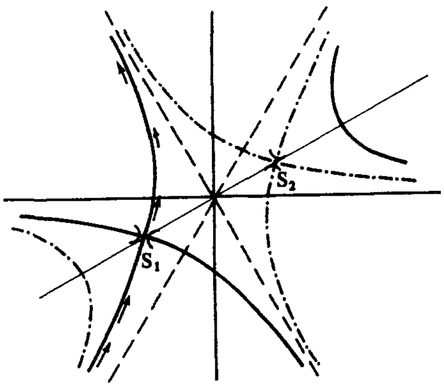


Fig. 9.8. $\frac{1}{3}\pi < \arg\zeta < \frac{2}{3}\pi$. See captions for figs. 9.5, 9.6.

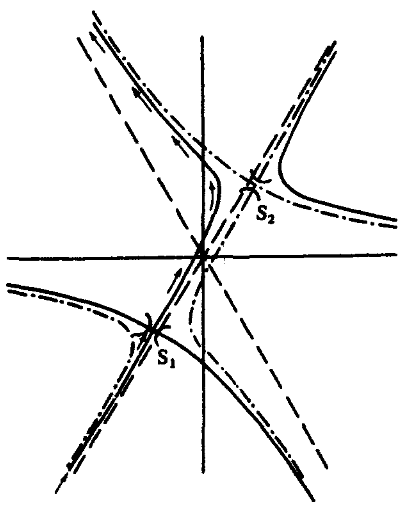


fig. 9.8 the path approaches the saddle point S_2 at $t_0 = me^{\frac{1}{3}i\theta}$, but turns fairly sharply through nearly a right angle to the left, and ends in the correct sector.

Now consider the curves shown in fig. 9.9 for $\theta = \frac{2}{3}\pi$ which is a Stokes line. They are like those for $\theta = 0$, but rotated through 60°. The line of steepest descents through the saddle point S_1 at $t_0 = -me^{\frac{1}{3}i\theta}$ is a straight line which begins in the correct sector (at $\arg t = -\frac{2}{3}\pi$), ascends to the saddle point S_1 and then descends and reaches the other saddle point S_2. To continue the descent, it is necessary to turn through a right angle either to the right or to the left. For the integration path we choose a left turn as indicated by the arrows, so that the path ends in the correct sector ($\arg t = \frac{2}{3}\pi$), and the approximate value of the integral is (9.42) as before.

Fig. 9.9. $\operatorname{Arg} \zeta = \frac{2}{3}\pi$. Stokes line.

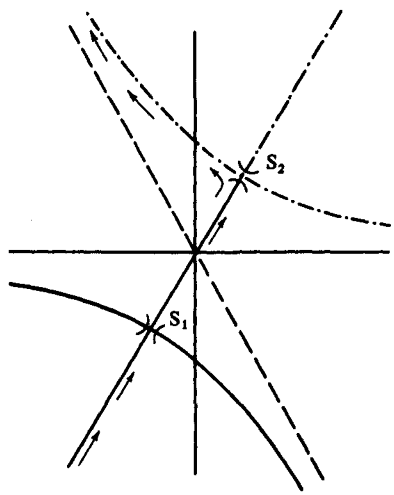


Fig. 9.10. $\frac{2}{3}\pi < \arg \zeta < \pi$. See captions for figs. 9.5, 9.6.

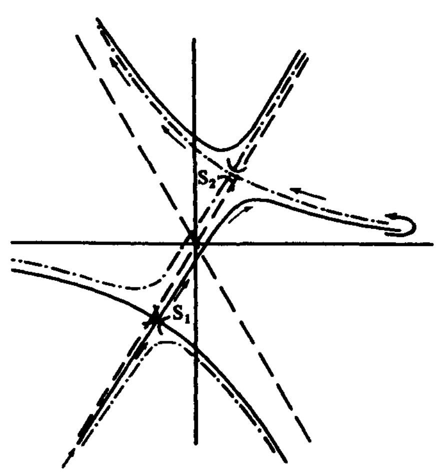


Fig. 9.10 shows the curves for $\frac{2}{3}\pi < \theta < \pi$. The line of steepest descents through the saddle point begins in the correct sector, passes through the saddle point, but when it approaches the other saddle point S_2 it now turns fairly sharply through a right angle to the right, and ends near the asymptote $v = 0$, which is not in the correct sector for the end of the path. The line of steepest descents through the saddle point S_2, however, begins near the asymptote $v = 0$, and ends where $\arg t = \frac{2}{3}\pi$. The path is therefore distorted so as to coincide with both lines of steepest descent, as shown by the arrows in fig. 9.10, and, when the integral is evaluated, the contributions from both saddle points must be included.

The saddle point S_2 is where $t_0 = m\mathrm{e}^{\frac{1}{2}\mathrm{i}\theta}$, and equations (9.37) and (9.38) show that $f_0 = \frac{2}{3}\zeta^{\frac{3}{2}}$ (where $\arg\ \zeta^{\frac{3}{2}}$ is $\frac{3}{2}\theta$) and $f_0'' = -2m\mathrm{e}^{\frac{1}{2}\mathrm{i}\theta}$, so that $\alpha = \pi + \frac{1}{2}\theta$, $A = 2m$. The direction Θ of the contour at the saddle point is $\pi - \frac{1}{4}\theta$, so that the minus sign is used in (9.29). It gives for the approximate contribution to the integral at the saddle point S_2:

$$\tfrac{1}{2}\pi^{-\frac{1}{2}}|m^{-\frac{1}{4}}|\exp\{\mathrm{i}(\tfrac{1}{2}\pi - \tfrac{1}{4}\theta)\}\exp(\tfrac{2}{3}\zeta^{\frac{3}{2}}). \qquad (9.44)$$

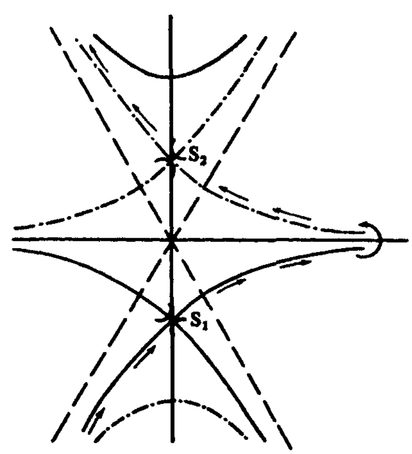
The contribution from the saddle point S_1 is (9.42) as before, and the two terms together give

$$\mathrm{Ai}(\zeta) \sim \tfrac{1}{2}\pi^{-\frac{1}{2}}\zeta^{-\frac{1}{4}}\{\exp(-\tfrac{2}{3}\zeta^{\frac{3}{2}}) + \mathrm{i}\exp(\tfrac{2}{3}\zeta^{\frac{3}{2}})\}, \quad (\tfrac{2}{3}\pi \leqslant \arg\zeta \leqslant \tfrac{4}{3}\pi) \qquad (9.45)$$

where $\arg(\zeta^{\frac{3}{2}}) = \frac{3}{2}\theta$; $\arg(\zeta^{-\frac{1}{4}}) = -\frac{1}{4}\theta$. The factor i in (9.45) is the Stokes multiplier.

The above argument shows how the second term comes in discontinuously at the Stokes line $\theta = \frac{2}{3}\pi$. When $\theta - \frac{2}{3}\pi$ is positive but small, the value of the integrand $\exp(\zeta t - \frac{1}{3}t^3)$ is much smaller at S_1 than at S_2, and the second term in (9.45) is small; it is the subdominant term. As $\arg\zeta$ approaches π, however, the two terms become more nearly equal, and on the anti-Stokes line at $\theta = \pi$ they have equal moduli. The curves for $\theta = \pi$ are shown in fig. 9.11. They are like those for $\theta = \frac{1}{3}\pi$, but rotated

Fig. 9.11. Arg $\zeta = \pi$. Anti-Stokes line.



through 60°. Here ζ is real and negative, and the two saddle points are on the imaginary axis. The curves could be traced in a similar way for other values of θ. For the Stokes line at $\theta = \frac{4}{3}\pi$ they would be like those for $\theta = \frac{2}{3}\pi$, but rotated through 60°. For $\frac{4}{3}\pi < \theta < 2\pi$ it is again possible to choose the contour so that it coincides with a line of steepest descents through only one saddle point, and in this range there is only one term in the asymptotic approximation. When $\theta = 2\pi$, the curves again become as in fig. 9.5. When θ increases from 0 to 2π, the two saddle points move only through an angle π in the complex t plane, so that their positions are interchanged. This is illustrated by the fact that the line in the Stokes diagram (fig. 8.6) for $\mathrm{Ai}(\zeta)$ is not a closed curve.

9.8. The method of stationary phase

Instead of using the method of steepest descents, it is possible to evaluate integrals of the type (9.1) by the method of stationary phase. In this method the path is distorted so as to pass through the same saddle points as for the method of steepest descents, but instead of following the lines of steepest descents, it is made to coincide with the level lines through the saddle points. The contributions from the neighbourhood of each saddle point used are then evaluated and added.

Whenever the path passes through a saddle point it must run anywhere within the two valleys on each side. For those parts of a valley which are very distant from the saddle point, the valley has steep walls, as explained in § 9.4, and a level line through the saddle point runs in one of these walls. The path can be moved up to either of these level lines. At a place remote from all saddle points, the steep wall of a valley may contain the level lines from several different saddle points, and their separation in the complex t plane is small because of the steepness of the wall. If, therefore, more than one saddle point is used, the level lines through them must be linked together, and this is done at points in steep hillsides remote from the saddle points, where the lengths of the links are extremely small, so that their contributions to the integral are negligible. The choice of the path for the method of stationary phase usually requires considerably more care than for steepest descents, partly because in any one valley there is only one line of steepest descent through the saddle point, whereas there are two level lines through it.

A level line may be chosen to run continuously through a saddle point or to turn through a right angle there. It is now assumed to run continuously. The case where there is a right angle turn can be dealt with by a simple extension of the following method.

The contribution to (9.21) from near one saddle point, at $t = t_0$, is found as follows. The exponent $f(t)$ is expanded in a Taylor series about $t = t_0$, and, since $\mathrm{d}f/\mathrm{d}t$ is zero there, the result leads to (9.24) as before. On a level line through the saddle point the last exponent in curly brackets in (9.24) is purely imaginary, and is zero at the saddle point. There are two level lines through the saddle point, and for one the phase is a minimum and for the other it is a maximum. Suppose that we choose the one on which it is a minimum. Then we take

$$\frac{1}{2!}(t-t_0)^2 f_0'' + \frac{1}{3!}(t-t_0)^3 f_0''' + \cdots = \mathrm{i}\tau^2 \tag{9.46}$$

where τ is real. Now if $|t-t_0|$ is small enough the second and later terms of (9.46) are small compared with the first. Thus

$$\mathrm{d}t \approx \pm (f_0''/2)^{-\frac{1}{2}} \exp(\tfrac{1}{4}\mathrm{i}\pi)\mathrm{d}\tau. \tag{9.47}$$

The sign depends on the angle Θ between the path and the real t axis at the saddle point. It is positive if $-\frac{3}{4}\pi < \Theta \leqslant \frac{1}{4}\pi$. The integral of (9.24) thus contains the factor

$$\int_{-\infty}^{\infty} \exp(\mathrm{i}\tau^2)\,\mathrm{d}\tau = \pi^{\frac{1}{2}} \exp(\tfrac{1}{4}\mathrm{i}\pi) \tag{9.48}$$

from (9.19) so that the contribution to (9.21) becomes

$$\pm (2\pi/A)^{\frac{1}{2}} \exp(f_0) \exp\{\tfrac{1}{2}\mathrm{i}(\pi-\alpha)\} \tag{9.49}$$

where $f_0'' = A\mathrm{e}^{\mathrm{i}\alpha}$ and $0 \leqslant \alpha < 2\pi$. This is exactly the same as the result (9.29) obtained by the method of steepest descents. If the other level line through the saddle point is used, so that the phase is a maximum there, it can be shown that again the result is the same.

In the method of steepest descents the integral is reduced to the form in (9.29) which is the error integral (9.14), and the integrand is always real. In the method of stationary phase the integral is reduced to the complex Fresnel integral (9.48) (compare (9.19)), in which the integrand has modulus unity, but its argument or phase varies along the contour and is proportional to τ^2. Its value is thus the resultant of a Cornu spiral, fig. 9.4(b). The main contribution is from near the saddle point where small contributing vectors have nearly the same direction because the phase is stationary there. On receding from the saddle point the phase changes more and more rapidly, so that the curve spirals in more and more quickly.

A good illustration of the use of the method of stationary phase is provided by the contour integral (8.16) for $\mathrm{Ai}(\zeta)$ when ζ is real and negative. Then the saddle points $t_0 = \pm\zeta^{\frac{1}{2}}$ lie on the imaginary t axis, which is a level line through both saddle points. The contour may be distorted to coincide with the imaginary axis, and it is convenient to let $t = \mathrm{i}s$. This was done in § 8.6 and was shown to lead to Airy's expression (8.19). For the method of stationary phase, however, we use the form (8.18). The saddle points are where $s = \pm|\zeta^{\frac{1}{2}}|$. When $s = +|\zeta^{\frac{1}{2}}|$ the phase $\zeta s + \frac{1}{3}s^3$ is a minimum, and the Cornu spiral is as in fig. 9.4(b). When $s = -|\zeta^{\frac{1}{2}}|$ the phase is a maximum, and the Cornu spiral has the opposite curvature to that in fig. 9.4(b). The asymptotic approximation to $\mathrm{Ai}(\zeta)$ is therefore the sum of the resultants of two equal Cornu spirals with opposite curvature. This was illustrated in more detail in § 8.22 and fig. 8.13.

The contributions from the two saddle points may each be expressed in the form (9.49). It can be shown that when they are added together with the correct signs, they lead to the same expression (9.45) as was obtained by the method of steepest descents.

9.9. Higher order approximation in steepest descents

In the method of steepest descents, as described in § 9.6, it was assumed that the second and later terms of the series (9.25) could be neglected, and this led to the approximation (9.26) and to the approximate value (9.29) for the contribution to the integral. We must now examine what happens when the approximation is not made. This leads to a more accurate value for the contribution (9.29) or (9.31) or (9.33) and to a quantitative criterion to show when the approximation (9.26) can be used.

The change of variable (9.25) can be written

$$s^2 = f_0 - f(t) \tag{9.50}$$

and then the integral (9.1) becomes

$$I = \exp(f_0)\int_{-\infty}^{\infty} \exp(-s^2)\phi(s)\,\mathrm{d}s \tag{9.51}$$

where

$$\phi(s) = g(t)\,\mathrm{d}t/\mathrm{d}s. \tag{9.52}$$

From (9.25), s can be expressed as a series in powers of $t - t_0$, and then, by reversion (Abramowitz and Stegun, 1965, p. 16), $t - t_0$ can be expressed as a series in powers of s. The first term is found from (9.25) simply by neglecting the third and higher powers of $t - t_0$. Hence

$$t - t_0 = \pm \mathrm{i}s\sqrt{2}(f_0'')^{-\frac{1}{2}}\{1 + \tfrac{1}{2}C_1 s + \tfrac{1}{3}C_2 s^2 + \cdots\}. \tag{9.53}$$

The sign depends on the direction Θ of the path in the t plane at the saddle point. It is $+$ if $-\frac{1}{2}\pi < \Theta \leqslant \frac{1}{2}\pi$. The $+$ sign is used here. Then it can be shown that

$$C_1 = \frac{\mathrm{i}\sqrt{2}}{3} f_0'''(f_0'')^{-\frac{3}{2}}$$

$$C_2 = \{\tfrac{1}{4}f_0''''f_0'' - \tfrac{5}{12}(f_0''')^2\}(f_0'')^{-3}. \tag{9.54}$$

Now suppose, first, that only the first term of (9.25) is appreciable. Then $f(t) - f_0$ is proportional to $(t - t_0)^2$ so that $\exp\{f(t)\}$ has only one saddle point, and no singularities except at infinity. We may say that its saddle point is 'isolated'. We also assume that $g(t)$ is slowly varying and can be taken as a constant $g(t_0)$, as was done in (9.31). Then (9.52), (9.53) give

$$\phi(s) = \pm\,\mathrm{i}\{\tfrac{1}{2}f_0''\}^{-\frac{1}{2}}g(t_0) \tag{9.55}$$

which is a constant, and was used to get the simple first order result (9.31).

Next suppose that the first two terms of (9.25) are appreciable but the others are negligible. Now $f(t)$ has two saddle points, one where $t = t_0$ and the other where

$$t = t_0 - 2f''_0/f'''_0. \tag{9.56}$$

If f'''_0 is very small, the saddle points are a long way apart, and we may still say that the one at t_0 is 'isolated'. Then (9.55) is still a good approximation for $\phi(s)$ and we can expect to use (9.31) as before. But if f'''_0 is not small, $\phi(s)$ is not a constant and we must use more than just the first term of (9.53).

If more than two terms of (9.25) are appreciable, $f(t)$ has more than two saddle points. If the coefficients of the second and third terms are small enough, the other saddle points are a long way from t_0.

Now from (9.52), (9.50)

$$\phi(s) = -2sg(t)/f'(t). \tag{9.57}$$

But $f'(t)$ is zero at a saddle point. For $t = t_0$, s is also zero and ϕ is bounded. But for any other saddle point ϕ is infinite, so that the series (9.53) cannot converge. Hence we have the very important result: for all functions $f(t)$ except a simple quadratic function, the series (9.53) cannot converge for all values of s. It must have a finite radius of convergence R. This is always less than the distance from t_0 to the next nearest saddle point, or to the next nearest singularity of $f(t)$.

If the saddle point at t_0 is sufficiently 'isolated', the radius of convergence R may be large, and a few of the coefficients $C_1, C_2 \cdots$ after the first in (9.53) may be small. But the series must ultimately diverge where $|s| > R$. In spite of this divergence of (9.53), we insert its derivative with respect to s in (9.52), (9.51), and integrate term by term. The odd powers of s then give zero, but the even powers give

$$I \sim \exp(f_0)\, P\pi^{\frac{1}{2}}(1 + \tfrac{1}{2}C_2 + \tfrac{3}{4}C_4 + \cdots) \tag{9.58}$$

where P is given by (9.28), (9.29). This series is the contribution to I from near the saddle point at t_0. It also is divergent, but nevertheless is useful when the first few terms after the first get successively smaller. It can be shown that it is an asymptotic series in the sense of Poincaré; see §8.11. The proof of this is known as Watson's lemma. It is a difficult subject and beyond the scope of this book.

The distance from t_0 to the next nearest saddle point depends on some parameter in the integral (9.1) or (9.51). For example in the Airy integral (8.16), the separation of the two saddle points is $2|\zeta^{\frac{1}{2}}|$, and is large when ζ is large. Here ζ is the parameter. When $|\zeta|$ is large the two saddle points are isolated and the simple form of the method of steepest descents should have good accuracy.

For this case

$$f(t) = \zeta t - \tfrac{1}{3}t^3 \tag{9.59}$$

whence it can be shown, from (9.54), that

$$\tfrac{1}{2}C_2 = -\tfrac{5}{48}\zeta^{-\frac{3}{2}} \tag{9.60}$$

and this is the same as the second term of the asymptotic series in (8.30), as given by (8.32), with $m = 1$. When this term is $\ll 1$, it is sufficiently accurate to use only the first order steepest descents formula (9.29) at each contributing saddle point. If it is required that the modulus of (9.60) shall be less than about 1/10, then the condition is $|\zeta| \geqslant 1$. We have already seen, § 8.10, equation (8.29), that this is also the condition that the W.K.B. solutions (8.27), (8.28) shall be good approximations to solutions of the Stokes equation.

The derivation of the series (9.58) via (9.54) is extremely cumbersome algebraically and this method is rarely used in actual computations. When an asymptotic series is needed it is usually easier to derive it in some other way, such as by the method of §8.11. The main use of C_2 in (9.54) is to test the accuracy of the first order steepest descents result (9.29) as illustrated by the example of (9.59), (9.60).

When two saddle points of (9.1) are close together but remote from all other saddle points and singularities, there is another method for finding their combined contribution. In their neighbourhood the exponent $f(t)$ cannot be treated as a quadratic function as was done in (9.26). Instead it is equated to a cubic function $\zeta s - \frac{1}{3}s^3$ which replaces the $-s^2$ in (9.25). It is arranged that the two saddle points where $\mathrm{d}f/\mathrm{d}t$ is zero in the t plane map into the two turning points $s = \pm\zeta^{\frac{1}{2}}$ of the cubic. Then, instead of the error integral in (9.29), an Airy integral appears. The resulting contribution is thus expressed in terms of Airy integral functions. In this way an analogue of the method of steepest descents can be constructed for two saddle points. It can be carried to higher orders of approximation in analogy with (9.58). It is then found that there are contributions from $\mathrm{Ai}(\zeta)$ and $\mathrm{Ai}'(\zeta)$, each multiplied by a series as in (9.58). The full theory was given in a most important paper by Chester, Friedman and Ursell (1957). The method has been used in a radio propagation problem (not described in this book) by Al'pert, Budden, Moiseyev and Stott (1983). An extension of the method is mentioned in § 10.20.

9.10. Double steepest descents

In wave propagation problems it is often necessary to evaluate double integrals of the form

$$\iint g(t, u) \exp\{f(t, u)\} \mathrm{d}t\, \mathrm{d}u \tag{9.61}$$

where t and u are complex variables. See, for example, (10.2). The methods of the earlier sections may be applied to the two variables t and u in turn. Suppose that the t integration is done first. Then u is held constant, the saddle points t_s where

$$\partial f/\partial t = 0 \tag{9.62}$$

are found, and the formula (9.31) is used to give the first order steepest descents contribution at each saddle point. It contains the factor $(\partial^2 f/\partial t^2)^{-\frac{1}{2}}$ as shown by

(9.27). From the physical conditions it is usually easy to decide which saddle points contribute to the integral. Quite often there is only one. In this integration it is assumed that $g(t, u)$ is slowly varying near the saddle point, so that it is treated as a constant $g(t_s, u)$; see § 9.6.

Next the u integration is done. The saddle points in the u plane are where $\mathrm{d}f\{t_s(u), u\}/\mathrm{d}u = 0$. But (9.62) remains true so that this is the same as the two conditions

$$\partial f/\partial t = 0, \quad \partial f/\partial u = 0. \tag{9.63}$$

These give a double saddle point $t = t_s$, $u = u_s$ in the complex t and u planes. The formula (9.31), with u for t, is again applied, and again the factors other than the exponential are treated as constants. Then it can be shown that the first order steepest descents contribution to (9.61) from the double saddle point is

$$\pm 2\pi g(t_s, u_s)\Delta^{-\frac{1}{2}}\exp\{f(t_s, u_s)\} \tag{9.64}$$

where

$$\Delta = \frac{\partial^2 f}{\partial t^2}\frac{\partial^2 f}{\partial u^2} - \left(\frac{\partial^2 f}{\partial t\, \partial u}\right)^2 \text{ at } t = t_s,\ u = u_s. \tag{9.65}$$

The important feature of the result (9.64) is that the exponential in (9.61) simply takes its value at the double saddle point. The paths of integration in the t and u planes can be thought of as lines of steepest descents through the double saddle point, but in physical problems it is often more convenient to take them as level lines. The condition (9.63) then shows that the phase of $\exp\{f(t, u)\}$ is stationary with respect to t and u at the double saddle point.

Integrals of the type (9.61) occur in §§ 10.2, 10.18, 10.19, 11.14. The use of double steepest descents in radio propagation problems has been discussed by Budden and Terry (1971), Budden (1976).

PROBLEMS 9

9.1. Evaluate $\int_{-\infty}^{\infty}\exp\{a(1-\cosh x)\}\mathrm{d}x$ approximately by the two following methods, given that $\mathrm{Re}(a) > 0$. (a) Expand the exponent in powers of x and ignore powers greater than x^2. (b) Put $a(1-\cosh x) = -u^2$. Express $\mathrm{d}x/\mathrm{d}u$ as a series of ascending powers of u and integrate two or three terms. Hence show that in method (a) the fractional error is about $1/8a$, that is, very small when $|a|$ is large.

Try the same methods for $\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi}\exp\{a(1-\sec\theta)\}\,\mathrm{d}\theta$.

9.2. Evaluate $\int_{-\frac{1}{2}\pi}^{\frac{1}{2}\pi}\exp(-a\tan^2 x)\mathrm{d}x$ as an asymptotic expansion given that $\mathrm{Re}(a) > 0$.

9.3. The Hankel function $H_0^{(1)}(z)$ is given by the Sommerfeld contour integral $(1/\pi)\int_{\mathrm{C}}\exp(\mathrm{i}z\cos u)\mathrm{d}u$ where the path C runs from $u = -\frac{1}{2}\pi + \mathrm{i}\infty$ to $u = \frac{1}{2}\pi - \mathrm{i}\infty$.

Evaluate the integral approximately by the method of steepest descents and hence show that the asymptotic form for $H_0^{(1)}(z)$ is $(2/\pi z)^{\frac{1}{2}} \exp\{\mathrm{i}(z-\frac{1}{4}\pi)\}$.
(See Sommerfeld, 1949, § 19).

9.4. The factorial function is defined by the integral $z! = \int_0^\infty t^z \mathrm{e}^{-t} \mathrm{d}t$. Use the method of steepest descents to show that when $|z|$ is large an approximate value of $z!$ is $(2\pi z)^{\frac{1}{2}} z^z \mathrm{e}^{-z}$ (first term of Stirling's formula). Sketch the paths used in the method of steepest descents when arg z is (a) zero, (b) $\frac{1}{2}\pi$, (c) $\pi - \varepsilon$ where ε is small.

9.5. The function $\exp\{f(t)\}$ has a double saddle point at $t = t_0$ (where both $\mathrm{d}f/\mathrm{d}t = 0$ and $\mathrm{d}^2 f/\mathrm{d}t^2 = 0$). Show that through this point there are six level lines (three crossing lines) with angles $\frac{1}{3}\pi$ between adjacent lines. Show also that there are three lines of steepest descent and three lines of steepest ascent, forming three crossing lines, and that these are at angles $\frac{1}{6}\pi$ to the level lines. Sketch the contour map for $\exp\{f(t)\}$ near t_0.

9.6. Use the method of steepest descents to find an approximation, valid when $a \gg 1$, for $F(a,m) = \int_0^\infty \exp(ax^m - x)\,\mathrm{d}x$ where m is a real constant, and $0 < m < 1$. In particular show that $F(a,\frac{1}{2}) \sim a\pi^{\frac{1}{2}} \exp(\frac{1}{4}a^2)$.
(Natural Sciences Tripos 1971. Part II. Theoretical Physics).