Qualitative Asymptotic Analysis of Complex Functions

Kenneth Man-kam Yip*
Department of Computer Science
Yale University
P.O. Box 2158, Yale Station
New Haven, CT 06520.
yip-ken@cs.yale.edu

Abstract

Professional scientists and engineers frequently analyze problems with heuristic techniques to get crude solutions quickly without worrying about the details of mathematical justification and firm error estimates. The method of steepest descent is one such technique that is widely used to evaluate functions represented by complex integrals. Routine application of the method however requires the selection of an appropriate integration contour in the complex plane. The selection task, which is difficult to do analytically, is often accomplished by geometric reasoning about the singular points of the integrand. This paper presents an implemented program that captures this style of reasoning. The program is based on a contour selection heuristic that reformulates the task as a simple graph search. Combined with symbolic algebraic techniques, the contour selection heuristic allows the program to quickly find the leading term approximation to the integral when a parameter in the integral becomes large.

Introduction

Finding out which qualitative reasoning style is good for which application is an important research topic in Qualitative Reasoning. Much of the previous qualitative reasoning work deals with mathematical functions specified by qualitative, algebraic, difference, or differential equations. However, an extremely important class of useful functions, including but not limited to the so-called *special functions* of mathematical physics, has integral representations. These functions are often solutions of certain frequently occurring linear second order differential equations *with variable coefficients* which cannot be solved in closed forms in terms of the elementary functions. Nevertheless it is sometimes possible to find a representation of the solution in terms of an integral with the independent variable appearing as a parameter. Except in rare cases, these integrals cannot be integrated in close form. Fortunately simplification of the integral is often possible when a parameter in the integral becomes large.

*Supported in part by NSF Grant CCR-9109567.

Asymptotic analysis of integrals is the branch of mathematics which is concerned with techniques to obtain approximate analytical representations of the behavior of integrals in the limit of a large parameter. The objective of asymptotic analysis is to find simple approximate representations of complex mathematical objects. For example, saying that the solutions to a differential equation are expressible in terms of modified Bessel functions of order $\frac{1}{6}$ does not convey much qualitative information to someone who is not already an expert in special functions. However, a simpler representation of the solutions in terms of $\frac{1}{x}e^{\pm\frac{x^3}{3}}$ as $x \to \infty$ is much more useful for deducing qualitative behaviors and for comparison with experimental or numerical data.

We study integrals not only because they arise in diverse areas of application, but also because the analysis of integrals and analysis of differential equations are both necessary parts in the study of mathematical models. There are important mathematical functions, such as the Mathieu functions, which cannot be represented by integrals. On the other hand, the factorial function, which has an integral representation, satisfies no differential equations of finite order [Jeffreys, 1962]. Neither integrals nor differential equations can replace each other.

To motivate integral representations, we use four examples to illustrate how these integrals arise in the study of physical problems.

Example 1. The solution to a linear differential equation, ODE or PDE, can often by written as an integral via the inverse Laplace or Fourier Transforms [Zwillinger, 1992]. For example, consider the initial value problem

$$ay'' + by' + cy = f(t)$$

with $y(0) = y_0$ and $y'(0) = y'_0$. The solution can be expressed as the inverse Laplace transform of some function $Y(s)$ involving the Laplace transforms of $f(t)$, the coefficients, and the initial values:

$$y(t) = L^{-1}(Y(s)) = \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} Y(s)e^{st}ds$$

where $\sigma > 0$ lies to the right of all the singularities of the integrand $Y(s)$ in the complex plane. Asymptotic analysis can often give a simple formula describing the large-t behavior of the solution.

Example 2. Solutions to linearized problems in wave propagation (such as radio or water or quantum-mechanical waves) are typically represented by superposition of simple progressive waves of the form $e^{i(kx-\omega t)}$ where k is the wave number and ω is the frequency. A formal solution could be obtained by summing over all wavenumbers:

$$\eta(x,t) = \int_{-\infty}^{\infty} A(k)e^{i(kx-\omega t)}dk$$

where $A(k)$ is related to initial conditions. Such formal solution is in general difficult to analyze analytically or numerically. Simplification of the integral can often be obtained if we are interested in the long time or farfield behaviors of the integral.

Example 3. The Gamma Function, denoted by$\Gamma(z)$, is a complex-valued generalization of the factorial function; it finds applications in number theory and approximation theory.

$$\Gamma(z+1) = \int_0^{\infty} t^z e^{-t}dt$$

which is valid for $Re(z) > -1$ with the property that $\Gamma(n+1) = n!$ for any non-negative integer n. For large n, the Stirling formula $n! \sim \sqrt{2\pi}n^{n+\frac{1}{2}}e^{-n}$ describes the essential behavior of the factorial function.

Example 4. The probability distribution function of the sum of identically distributed random variables $X_1, X_2, \ldots, X_n$ can be expressed via the inverse Fourier transform as:

$$f_n(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \phi(\omega)^n e^{-i\omega t}d\omega$$

where $\phi(\omega)$ is the Fourier transform of the probability density function of each X_i. Determining the behavior of the integral as $n \to \infty$ is essentially the content of the fundamental central limit theorem in probability.

The purpose of this paper is to demonstrate how a particular asymptotic technique, the method of steepest descent [Bender and Orszag, 1978], can be automated to produce approximate analytical representations for a wide class of integrals. The technique is selected for three reasons: (1) it is widely used in many branches of physics, (2) it shares the same basic idea with a family of asymptotic techniques (such as the Method of stationary phase, and Laplace's Method), and (3) it gives higher order approximations. The method however is extremely tedious to apply even with the help of a conventional computer algebra system because it involves geometric reasoning about locations of the singularities and topography of the integrand in the complex plane. This paper shows how this problem can be solved by combining well-known geometric, symbolic, and numerical techniques.

Our work is rooted in the tradition of focusing on the problem-solving behavior of articulate professionals in well-structured domains and formalizing their methods so that a computer can exhibit similar behavior on similar problems. Previous works with a similar intent include [Abelson *et al.*, 1989; Sacks, 1990; Sacks, 1991; Yip, 1991; Zhao, 1991]. These authors dealt with problems specified by low-dimensional ordinary differential equations. Our project differs from these works in two major aspects. First, it is the first attempt to automate asymptotic analysis techniques for integrals. Second, whereas previous analysis employs representations based on signs of quantities or phase space descriptions, we use analytical representations in terms of elementary functions, which are harder to obtain but are much more informative.

The Task

We are interested in the task of analyzing the asymptotic behavior of functions defined by complex integrals. The analysis is part of a much more involved process of modeling-analysis-verification typical in the study of physical phenomena. The analysis program takes three inputs: (1) a complex integral with a parameter, (2) the contour for integration, and (3) a description of the parameter. Its output is a simple formula (in terms of elementary functions) and its qualitative interpretation representing the asymptotic behavior of the function.

The integral may come from solving linear differential equations say by methods like Fourier Transform, Laplace Transform, or contour integration [Zwillinger, 1992]. The results of the asymptotic analysis are typically used to make predictions about the behavior of the system in question; such predictions are further compared with detailed numerical simulation or experimental data.

The analysis program works for a class of integrals of the following form:

$$I(\lambda) = \int_C g(z)e^{\lambda h(z)}dz \qquad (1)$$

where C is a contour in the complex plane, and $g(z)$, and $h(z)$ are analytic functions independent of λ. The parameter λ is assumed to be real, but there is no loss of generality because we can always include the phase factor of a complex parameter into the function $h(z)$.

As our model problem, we use the Airy function (or the rainbow integral), in its complex form:

$$Ai(\lambda) = \frac{1}{2\pi i}\int_C e^{\lambda z - \frac{1}{3}z^3}dz$$

which has found applications in many branches of physics – ray diffraction in optics, tunneling of quantum particles, and evolution of wavefront of tsunamis, just to name a few. The integral can also be expressed as a solution to the linear differential equation:

$$\frac{d^2w(\lambda)}{d\lambda^2} - \lambda w(\lambda) = 0 \tag{2}$$

The Airy equation has the same form as the time-independent Schrodinger equation in the neighborhood of a classical turning point.

The most difficult part of the analysis problem is the deformation of the given contour C to a new contour C' to which one can easily apply the method of steepest descent (see section 3). Suppose the original contour C is from $\infty e^{i\frac{4\pi}{3}}$ to $\infty e^{i\frac{2\pi}{3}}$ (see Fig. 4a). For positive λ, the program chooses a new contour passing through the point s_1 in the complex plane; for negative λ, it chooses a new contour consisting of two segments: one passing through s_1, and the other through s_2 (see Fig. 4c).

It is interesting to quote how an author of a textbook on the subject describes this difficulty.

> Any special application of the steepest descent method consists of two stages.
>
> (i) The stage of exploring, conjecturing, and scheming, which is usually the most difficult one. It results in choosing a new integration path, made ready for application of (ii).
>
> (ii) The stage of carrying out the method. Once the path has been suitably chosen, this second stage is, as a rule, rather a matter of routine... (From [de Bruijn, 1981, page 77].)

From the Airy equation (2), for $\lambda > 0$, heuristically we would expect the solutions to display some kind of exponential behavior; on the other hand, for $\lambda < 0$, we expect oscillatory behaviors. The output of the analysis program in Fig. 1 confirms this expectation.

For $\lambda \rightarrow \infty$:

$$Ai(\lambda) \sim \frac{1}{2}\frac{1}{\sqrt{\pi}}\frac{1}{\lambda^{\frac{1}{4}}}e^{-\frac{2}{3}\lambda^{\frac{3}{2}}}$$

The function decays exponentially with an exponent proportional to $\lambda^{\frac{3}{2}}$.
For $\lambda \rightarrow -\infty$:

$$Ai(\lambda) \sim \frac{1}{\sqrt{\pi}}\frac{1}{|\lambda|^{\frac{1}{4}}}\sin(\tfrac{2}{3}|\lambda|^{\frac{3}{2}} + \frac{\pi}{4})$$

The function oscillates with its amplitude decaying algebraically with a factor proportional to $\lambda^{-\frac{1}{4}}$.

Figure 1: *The program summarizes its findings in stylized English description. It has no natural language capability; the description is for us to read.*

Characteristics of the Problem Domain

Some Terminology

A complex variable z may be written as $z = x + iy$ where x and y are real. A complex function $h(z)$ can similarly be expressed in terms of its real and imaginary parts: $h(z) = u(x, y) + iv(x, y)$ where u and v are real-valued functions. A complex function $h(z)$ is **analytic** in a region R of the complex plane if it is differentiable at every point $z_0 \in R$. Given a point z_0, a directed curve from z_0 along which $u(x, y)$ is decreasing is called a **path of descent**. The tangent of path of descent at z_0 is called the **direction of descent**. In an analogous way, we define a **path of ascent** and **direction of ascent** for a directed curve from z_0 along which $u(x, y)$ is increasing. Emanating from z_0 there are many paths of descent and ascent. A **path of steepest descent** is a path of descent whose rate of descent is maximal; its direction is a **direction of steepest descent**. Completely analogous definitions can be given for a steepest ascent path and its direction.

A point z_1 is said to lie in the **valley** of $h(z)$ with respect to z_0 if $u(x_1, y_1) < u(x_0, y_0)$ and on the **hill** of $h(z)$ if $u(x_1, y_1) > u(x_0, y_0)$.

The lines for which $u(x, y)$ is constant is called the level lines for the function $e^{h(z)}$; on these level lines, $e^{h(z)}$ has constant magnitude because $| e^{h(z)} |=| e^{u+iv} |=| e^u | \times | e^{iv} |= e^u$. Since an analytic function satisfies the Cauchy-Riemann conditions, the level lines are orthogonal to the lines for which $v(x, y)$ is constant. Hence, the paths of steepest descent (ascents) coincide with the constant v-lines.

At point z_0 where $\frac{dh(z_0)}{dz} = 0$ the magnitude and phase of $e^{h(z)}$ is stationary. However, by the Maximum Modulus Theorem, u and v cannot have a maximum (or a minimum) inside the domain of analyticity of $h(z)$ unless $h(z)$ is identically constant [Carrier *et al.*, 1966]. The point z_0 is thus a **saddle point** of $h(z)$. A saddle point is **simple** if $h'(z_0) = 0$ but $h''(z_0) \neq 0$. (See Fig. 2.) A point is **regular** if it is not stationary.

Finally, two functions $f(z)$ and $g(z)$ are **asymptotically equivalent**, denoted by $f \sim g$, if $\lim_{x\rightarrow\infty} \frac{f(x)}{g(x)} = 1$.

Basic Idea of Approximation

The basic idea (usually attributed to Laplace) for evaluating integrals of the form $I(\lambda)$ is easy to understand if the integrand consists of real functions. The extension to complex functions will be dealt with in next section. Suppose the integrand is the product of two functions indicated in Fig. 3. The function $f(x)$ is narrow in the sense that it is sharply peaked within an interval on the x-axis which is small compared to the distances over which $g(x)$ changes significantly. Then it is plausible that the integral can be well approxi-

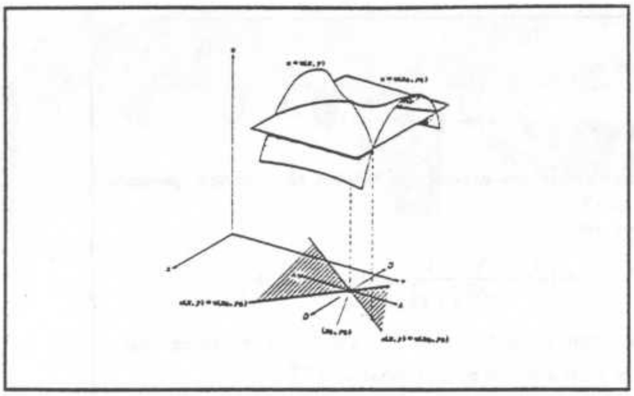

Figure 2: The surface of u(x,y) near a simple saddle point (x_0, y_0) The steepest descent paths are marked D, and the steepest ascent paths marked A. Note the alternating hills and valleys around the saddle point.

mated by:

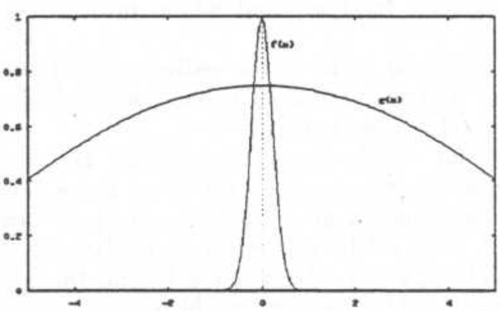

Figure 3: Laplace's idea of approximation. The function f(x) is sharply peaked, and g(x) can be well approximated by its value at the maximum of f(x).

$$\int_{-\infty}^{\infty} g(x)f(x)dx \approx g(x_0)A$$

where A is the area under $f(x)$, and x_0 is where $f(x)$ attains its maximum. To calculate A we only need to integrate in a small neighborhood around x_0 since $f(x)$ is close to zero outside the interval. Furthermore, the approximation will improve as $f(x)$ becomes more sharply peaked.

Laplace's method was developed for functions of the form $f(x) = e^{-\lambda h(x)}$. The *hat function* $e^{-\lambda x^2}$, for instance, has this "sharply-peaked" character. Since any analytic function $h(x)$ can be expanded in a Taylor series:

$$h(x) = h(x_0) + h'(x_0)(x - x_0) + \tfrac{1}{2}\, h''(x_0)(x - x_0)^2 + \dots$$

we expect, at a stationary point x_0, $h(x)$ would locally look like the hat function. With this observation, we can think of the purpose of deformation of the integration contour is to get to a stationary point where Laplace's method can be applied.

Method of Steepest Descent

A naive application of Laplace's idea to a complex integrand won't work. Suppose C is the original contour which starts and ends at a valley. Let the maximum of the real part of $h(z)$ on C be attained at a regular point z_0. A first guess of the order of estimate (big "oh") for the integral might be:

$$I(\lambda) = O(g(z_0)e^{\lambda u(z_0)}\Delta z_0)$$

where $u(z_0)$ is short for $u(x_0, y_0)$ and Δz_0 is the range over which $h(z)$ substains its maximum. But the guess is completely inaccurate because we have not considered the influence of the imaginary part $v(z_0)$. For large λ, $e^{i\lambda v(z_0)}$ oscillates rapidly and could cause complete cancellation in the integration so that the argument that the major contribution to the integral comes from the region of maximum is no longer valid.

So in order to apply Laplace's idea, we have to look for a contour where the imaginary part is constant (i.e., along the constant v-lines which, as we have seen, are just the paths of steepest descents or ascents). By Cauchy Integral Theorem, the contour C can be deformed without changing the value of $I(\lambda)$. Since we want the magnitude of $h(z)$ to decrease rapidly away from z_0, C should be deformed to the path of steepest descent emanating from z_0. But a regular point z_0 has only one path of steepest descent emanating from it. So for the contour to end in another valley, we must switch to another path of steepest descent in the neighborhood of z_0. But that means we can't stay on a single constant v-line and again we have the contribution from the $e^{i\lambda v}$ to contend with.

Now it is clear why the saddle point is important. A simple saddle point has *two* paths of steepest descent (with the same v value) emanating from it. Therefore, it seems that the saddle point and its steepest descent paths are the correct paths to apply the Laplace's method. The above informal reasoning can be backed up by formal arguments [Olver, 1974]. We will take the proof for granted and instead turn to the implementation of the method of steepest descent.

The method consists of four basic steps [Bleistein and Handelsman, 1986]:

1. Identify the critical points of the integrand of $I(\lambda)$, i.e., the endpoints of integration contour C and the saddle points of $h(z)$.
2. Determine the paths of steepest descent from each of the critical point.
3. Justify, via Cauchy's Integral Theorem, the deformation of the original contour C onto one or more of the paths of steepest descent found in step (2).
4. Sum the contribution to the integral from each of the paths of steepest descent. (Each integral is of the Laplace type.)

Steps 1,2 and 4 offer no conceptual difficulties – we will see how to do them in section 5. Step 3, the choice of a new contour, as alluded above is considered the hardest. The function $h(z)$ might have an infinite number of saddle points, but only a few of them will be relevant to the determination of $I(\lambda)$. A selection method based on conformal mapping is sometimes proposed [Olver, 1974], but it is too cumbersome to use. In the next section, we will see how an informal method can quickly find the desired contour.

Theory of Contour Selection

The topography of an analytic function is exceedingly simple: it can't have peaks or pits, and around any saddle point there are alternating hills and valleys. Furthermore, the valleys of two distinct saddle points can't overlap partially (again a consequence of the Maximum Modulus Theorem): either they are disjoint or one is completely included in the other. We will exploit these properties to come up with a simple graphical representation of the topography of $u(z)$, the real part of $h(z)$, and reduce the problem of contour selection to a shortest-path graph search.

In problems of physics, it is typically the case that the contour C begins and ends at valleys so that the integrand tends to zero as z tends to ∞. Informally, to deform the contour we find the valley where C begins and run the path over a saddle point along the steepest paths into another valley. If C ends at this valley, then the process is done; otherwise, run the path over another saddle point into a third valley. We repeat the process until we find the valley where C ends.

Since a valley may belong to more than one saddle point, we need a strategy to decide which saddle point to pick. Let's define some concepts. Given two simple saddle points s_1 and s_2, we say s_1 **dominates** s_2 if $s_2 \in V(s_1)$, where $V(s_1)$ is a valley of s_1. We say s_1 **immediately dominates** s_2 if there is no other saddle point s_3 such that s_1 dominates s_3 and s_3 dominates s_2. The saddle point s_1 **strictly dominates** s_2 if s_2 lies on a path of steepest descent emanating from s_1. The **admissible region** of $V(s_1)$ is defined to be:

$$AR(V(s_1)) = \begin{cases} V(s_1) & \text{if no other saddle point} \in V(s_1) \\ s_2 & \text{if } s_2 \in SDPATH(s_1, V(s_1)) \\ V(s_1) \cap AR(V(s_2)) & \text{if } SDPATH(s_1, V(s_1)) \cap V(s_2) \neq \phi \end{cases}$$

where $SDPATH(s_1, V(s_1))$ denotes the path of steepest descent from s_1 running down the valley $V(s_1)$.

Using these concepts, we can define a graphical representation for the topography of $u(z)$. The graph consists of two types of node and one type of edge. Each node represents an admissible region; it can be either a saddle point or a valley. Let's call the first type of node, S-node, and the second V-node. Each edge represents a steepest descent path from some saddle point. Two S-nodes are connected if one is on the steepest descent path of another. A S-node is connected to a V-node if the steepest descent path lies in the valley represented by the V-node. See Fig. 4a(ii). Now we can formulate our informal contour selection strategy more formally.

Contour Selection Heuristic: Given an integral $\int_C g(z)e^{\lambda h(z)}dz$ and a contour C with endpoints in val-

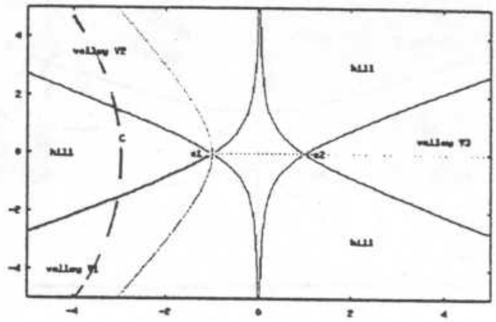

Figure 4: a(i): The topography of u(x,y) for $\lambda > 0$. The dotted lines are steepest descent paths through the saddle points s_1 and s_2. The original contour C marked by the dashed line is seen on the far left. The remaining solid lines are the boundaries of valleys and hills. Note that s_1 is strictly dominated by s_2.

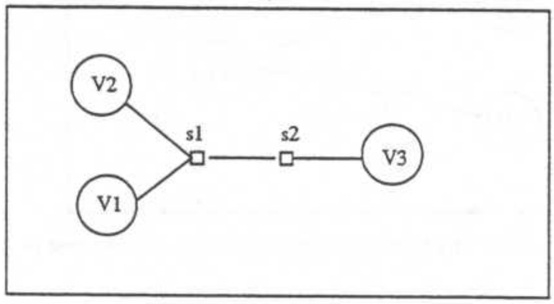

Figure 4: a(ii): The topography graph corresponding to that in Fig. 4a(i).

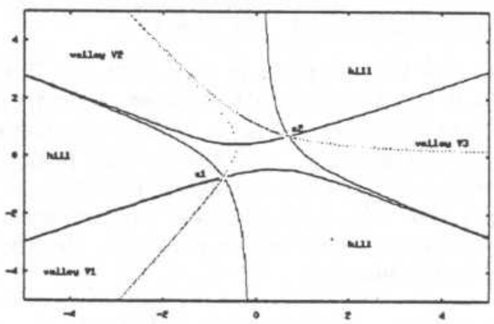

Figure 4: b(i): The topography of u(x,y) for $phase(\lambda) = \frac{\pi}{2}$. Note that s_2 is immediately dominated by s_1.

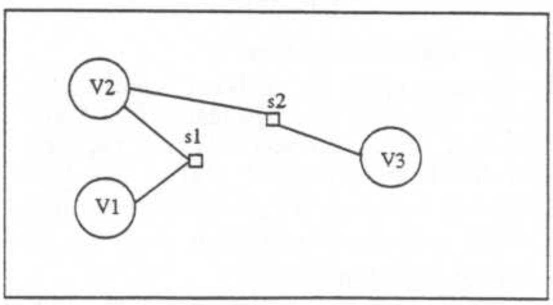

Figure 4: b(ii): The topography graph corresponding to that in Fig. 4b(i).

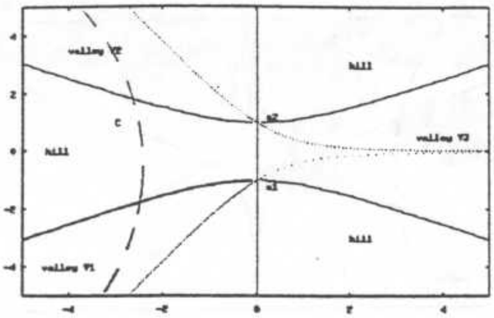

Figure 4: c(i): The topography of u(x,y) for $\lambda < 0$.

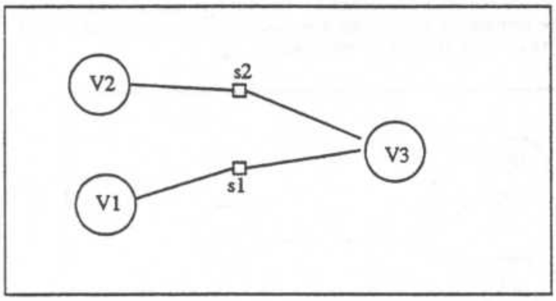

Figure 4: c(ii): The topography graph corresponding to that in Fig. 4c(i).
d

leys V_1 and V_2, deform it to a new contour represented by the shortest path connecting V_1 and V_2 in the topography graph of the real part of $h(z)$.

Let's see how this heuristic is used. For $\lambda > 0$, i.e., $phase(\lambda) = 0$, the shortest path between V_1 and V_2 is through the saddle point s_1. So s_1 is the only relevant saddle point. For $phase(\lambda) = \frac{\pi}{2}$ the shortest path still passes through only s_1 even the structure of the graph has changed. For $\lambda < 0$, the shortest path connecting V_1 and V_2 runs over both saddle points; so they both contribute to the integral.

Implementation: the details

In this section, we will describe how the four steps of the method of steepest descent (as mentioned in the end of section 3) are implemented.

Preliminary Transformation

Frequently the integral to be analyzed is not in the standard form where $h(z)$ is independent of λ. A preliminary rescaling of the variable z is thus required. For example, for the Airy integral, $h(z) = \lambda z - \frac{1}{3} z^3$. Solving $h'(z) = \lambda - z^2 = 0$, we get the saddle points $z = \pm\sqrt{\lambda}$. With a change of variable $w = \frac{z}{\sqrt{\lambda}}$, we get rid of the dependence of $h(z)$ on λ:

$$Ai(\lambda) = \frac{\sqrt{\lambda}}{2\pi i} \int_C e^{\lambda^{\frac{3}{2}}(w - \frac{1}{3}\, w^3)} dw$$

In general, the scaling factor is determined by the roots of $h'(z) = 0$. If the solution is $z = f(\lambda)$, then choose a new variable $w = z \times f(\lambda)$.

Finding Saddle Points

The saddle points are the solutions of the equation $h'(z) = 0$. Typically the function $h'(z)$ consists of simple polynomials in z, and functions of z such as sinh or cosh. We use the Mathematica `Solve` routine to find the roots. However, since Mathematica typically drops all but one root for multi-valued inverse functions, we need to patch it to recover all the roots. The capability of `Solve` is quite limited; for instance, it can't solve equations like $\log(1+z) = z$. So in the general case we will have to use numerical root-finding methods. Since root finding is not our research objective, we have not pursued this line of development.

Determining Valleys and Paths of Steepest Descent

This step is an important part of the algorithm. The valley boundaries of a saddle point are the constant u-lines emanating from it. To determine the boundary, we use a numerical continuation algorithm to track the constant u-lines starting from the saddle point. Tracking a curve given by an implicit equation $u(x, y) = c$, where c is a constant, can be tricky because the curve can hit a turning point, or a bifurcation point, where the curve is split into multiple branches. The algorithm makes uses of analytical information about the saddle point to guide the numerical continuation.

The procedure, `track-valley-boundaries`, takes four inputs: (1) an expression $h(z)$, (2) a saddle point s, (3) a stepsize, and (4) upper and lower bounds for x and y. Its output are the four constant u-lines emanating from s. The main steps of the algorithm are:

```
1. Compute u(x,y), the real part of h(z),
   symbolically.
2. Evaluate u(x,y) at s to get the height d
3. Compute the directions of steepest descent and
   ascent (by computing the second derivative of
   h(z) symbolically and evaluating it at s to
   get the phase). 1
4. Compute four starting points by taking a step
   of size h/2 in each of the four steepest
   directions.
5. Call continuation on each starting point.
```

The procedure, `continuation`, takes six inputs: (1) the lambda functions $u(x, y) - d$, $dudx(x, y)$, and $dudy(x, y)$, where $dudx$ and $dudy$ are the partial derivatives of u with respect to x and y respectively, (2) a start point (x_0, y_0), (3) an initial direction of the curve, (4) a stepsize h, maximum step size, and minimum step size, (5) upper and lower bounds for x and y,

[1] Let $re^{i\theta} = h''(s)$. Then the directions are given by $\frac{1}{2}\left(\frac{\pi}{2} - \theta\right), \pi + \frac{1}{2}\left(\frac{\pi}{2} - \theta\right), \frac{1}{2}\left(-\frac{\pi}{2} - \theta\right)$, and $\pi + \frac{1}{2}\left(-\frac{\pi}{2} - \theta\right)$. For details, see [Bleistein and Handelsman, 1986].

and (6) special terminating points. The special terminating points are the saddle points of $h(z)$; the continuation will stop if it runs into one of these points. The procedure returns a path on the constant u-line starting from (x_0, y_0). The algorithm has 4 steps:

```
1. Apply a corrector to the start point to get a
   more accurate initial point on the curve.
2. Apply a predictor to the current point to get
   a guess for the next point on the curve.
3. Apply a corrector to the predicted point.
4. Repeat steps 2 and 3 until the curve either
   exceeds the upper or lower bounds of x and y,
   or hits a special terminating point.
```

The predictor is based on arc-length continuation, a popular method to get past the turning point [Allgower and Georg, 1990]. Our implementation solves a set of two equations for two unknowns $\frac{dx}{dz}$ and $\frac{dy}{dz}$, where s is the arc length:

$$\frac{du}{ds} = \frac{\partial u}{\partial x}\frac{dx}{ds} + \frac{\partial u}{\partial y}\frac{dy}{ds} \tag{3}$$

$$(\frac{dx}{ds})^2 + (\frac{dy}{ds})^2 = 1 \tag{4}$$

Once the numerical values of $\frac{dx}{ds}$ and $\frac{dy}{ds}$ are found, an explicit multi-step 4th-order Adams-Bashforth is used to solve for the increments x and y [Kahaner *et al.*, 1989].

The corrector is Newton-Raphson with a small bound on the number of iterations allowed.

Finding the Steepest Descent Paths

Since the steepest descent (ascent) paths are identical with constant v-lines, we find them by tracking the function $v(x, y) = c'$, where c' is the value of $v(x, y)$ at the saddle point, using exactly the same continuation procedures for determining the valley boundaries. To tell which two of four v-lines are the descent paths, we just compare the value of $u(x, y)$ at each of the four starting points with that of the saddle point.

Constructing the Topography Graph

The construction of the graph depends on a number of well-known geometric algorithms. A valley is approximated by the convex hull of its valley boundaries. The convex hull is computed by the **Graham scan** algorithm, which runs in $O(n \log n)$ time where n is the number of points in S (see Shamos 1985). The output of the algorithm is the vertices of the convex hull arranged in the counterclockwise direction.

Testing if a saddle point s lies in the valley becomes a point-location problem for polygon. There are many algorithms to solve this problem. We simply compute the angles subtended by the edges of the convex hull about the point s. If s is inside the polygon, the total angle should sum to 2π.

An admissible region is approximated by intersection of convex polygons corresponding to the valleys. We use an algorithm invented by O'Rourke, which runs in $O(n + m)$ time where n and m are the number of vertices of the two convex polygons [Preparata and Shamos, 1985]. The output of the algorithm is another convex polygon ordered in the counterclockwise direction.

Finally, the strictly dominance relation is determined by the tracking algorithm when it reports hitting a saddle point.

As discussed in section 4, the desired contour is computed by a simple breadth-first search for the shortest path in the topography graph. The relevant saddle points, i.e., those that contribute to the integral, are those that lie on the shortest path.

Finding the Asymptotic Formula

Given the relevant saddle points, we symbolically expand $h(z)$ about them and truncate the series at the second derivative term to get a hat function as discussed in section 3.2. The Laplace type of integral is easily solved by Mathematica's `Integrate` command:

```
In[5]:= Sqrt[l]/(2 Pi I) * Integrate[Exp[- l^(3/2)(2/3 - x^2)],-x,-
Infinity,Infinity^]

                 -1
Out[5]= ---------------------
               3/2
         (2 l   )/3  1/4
       2 E          l    Sqrt[Pi]
```

Performance

Let us see in more detail the output of the analysis program when it is finding the valley boundaries and steepest descent paths. The procedure **display-landscape** takes three inputs: (1) the expression $h(z)$ in infix form, (2) the list of saddle points ($z = \pm 1$ in this run), and (3) a stepsize size (0.2 in this run). Notice that in tracking the second steepest descent path of the saddle point (1,0), the path hits the other saddle point (-1,0), causing a termination.

```
(display-landscape '(z - 1/3 * (z ^ 3)) '((1 0) ( -1 0)) .2))

Tracking valley boundaries of saddle point = (1.0000,0.0000)
Stepsize is 0.200
start point is: (1.0707,-0.0707) direction = (1.0,-1.0)
Total path points is: 33.
start point is: (1.0707,0.0707) direction = (1.0,1.0)
Stepsize increases to: 0.4000
Total path points is: 18.
start point is: (0.9293,0.0707) direction = (-1.0,1.0)
Total path points is: 33.
start point is: (0.9293,-0.0707) direction = (-1.0,-1.0)
Total path points is: 55.

Tracking its steepest descent curves...
Stepsize is 0.050
start point is: (1.0250,0.0000) direction = (1.0,0) descent angle = 0.0000
Total path points is: 70.
start point is: (0.9750,0.0000) direction = (-1.0,0) descent angle = 3.1416
Terminating at saddle point (-1.0000,0.0000)
Total path points is: 22.

Tracking valley boundaries of saddle point = (-1.0000,0.0000)
```

```
Stepsize is 0.200
start point is: (-0.9293,0.0707) direction = (1.0,1.0)
Stepsize increases to: 0.4000
Total path points is: 18.
start point is: (-1.0707,0.0707) direction = (-1.0,1.0)
Total path points is: 61.
start point is: (-1.0707,-0.0707) direction = (-1.0,-1.0)
Total path points is: 61.
start point is: (-0.9293,-0.0707) direction = (1.0,-1.0)
Total path points is: 55.

Tracking its steepest descent curves...
Stepsize is 0.050
start point is: (-1.0000,0.0250) direction = (0,1.0) descent angle = 1.5708
Total path points is: 130.
start point is: (-1.0000,-0.0250) direction = (0,-1.0) descent angle = -1.5708
Total path points is: 246.
```

A typical run takes less than 40 seconds real time on a Sparc 330. All the figures describing the topography of the real part of $h(z)$ in this paper are automatically produced by `display-landscape`.

Evaluation

We tested the program on the Airy function with several parameter values of λ (when its phase values are $0, \frac{\pi}{2}, \frac{2\pi}{3}$, and π), and on half a dozen of integrals, including the Gamma function, the Bessel function, and the Hankel function of the first kind. In each case, the program is able to reproduce the leading term of the asymptotic expansion as found in standard mathematical handbooks. In Fig. 5, we compare the exact Airy function with the approximate analytical expression obtained from asymptotic analysis for $\lambda < 0$. Notice the good agreement between the two graphs up to rather small values of λ (approx. -1) despite the fact the asymptotic formula is derived under the assumption of large negative λ.

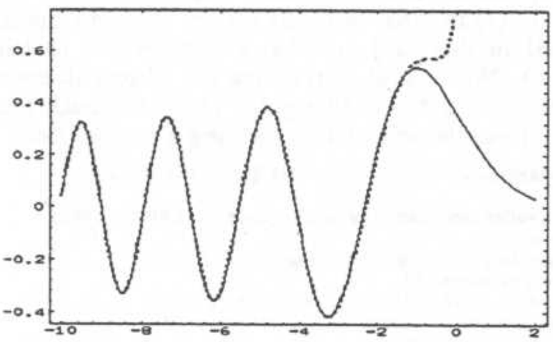

Figure 5: Comparison of the Airy function Ai(λ) (the solid line) and the function $\frac{1}{\sqrt{\pi}}\frac{1}{|\lambda|^{\frac{1}{4}}}\sin(\frac{2}{3} \mid \lambda \mid^{\frac{3}{2}} + \frac{\pi}{4})$ (the dotted line). Note the good agreement between the two graphs up to $\lambda < -1$.

Extension

The most obvious extension to the program is to allow the integrand to have singularities like poles, contours with finite end points, and branch cuts. Adding poles is a small task. It consists of detecting what poles are crossed over when the original contour is deformed, and summing the contribution of each pole according to the residual theory [Carrier *et al.*, 1966]. Detection can be reduced to the point location problem [Preparata and Shamos, 1985] if the the region bounded by the original and new contour is approximated by a convex polygon. Handling contours with finite end points is also easy – we add the end points to the topography graph with the paths of steepest descent from them. Unlike a saddle point, there is only one such path from each end point. Adding branch cuts requires some work because the program has to decide where to place the cuts. The integration around a branch cut is also slightly different from that around a saddle point. Another extension might be to generate higher order approximation beyond the leading term [Campbell *et al.*, 1987]. A third extension – being currently worked on – is to use the topography graph to automatically find all bifurcation behaviors of the integrals.

Conclusion

That the type of informal and heuristic reasoning for analyzing complex integrals as practiced by professional scientists can be captured in a computer program combining numerical, geometric, and symbolic techniques is the major contribution of this paper. The reformulation of the contour selection problem as a graph search allows a quick way to find the desired contour, a task usually considered difficult by textbook authors on the subject. Accomplished applied mathematicians [Carrier, 1972] consider the method of steepest descent an important technique that "has had a profound influence on the mathematical treatment of scientific and other real world problems." The present paper demonstrates that such powerful analysis technique can be brought within the grasp of computer programs today.

References

Abelson, H.; Eisenberg, M.; Halfant, M.; Katzenelson, M.; Sacks, E.; Sussman, G.; Wisdom, J.; and Yip, K. 1989. Intelligence in scientific computing. *Communications of the ACM* 32(5).

Allgower, Eugene and Georg, Kurt 1990. *Numerical Continuation Methods.* Springer-Verlag.

Bender, Carl and Orszag, Steven 1978. *Advanced Mathematical Methods for Scientists and Engineers.* McGraw-Hill.

Bleistein, Norman and Handelsman, Richard 1986. *Asymptotic Expansions of Integrals.* Dover.

Campbell, John; Froman, Per Olof; and Walles, Erik 1987. Explicit series formulae for the evaluation of integrals by the method of steepest descent. *Studies in Applied Mathematics* 77.

Carrier, George; Krook, Max; and Pearson, Carl 1966. *Functions of a Complex Variable: theory and technique.* McGraw-Hill.

Carrier, G.F. 1972. Heuristic reasoning in applied mathematics. *Quarterly Journal of Applied Mathematics.*

Bruijn, N.G.de 1981. *Asymptotic Methods in Analysis.* Dover.

Jeffreys, Harold 1962. *Asymptotic Approximations.* Oxford University Press.

Kahaner, David; Moler, Cleve; and Nash, Stephen 1989. *Numerical Methods and Software.* Prentice Hall. ISBN 0-13-627258-4.

Olver, F.W.J. 1974. *Asymptotics and Special Functions.* Academic Press.

Preparata, Franco and Shamos, Michael 1985. *Computational Geometry.* Springer-Verlag.

Sacks, Elisha P. 1990. Automatic qualitative analysis of dynamic systems using piecewise linear approximation. *Artificial Intelligence* 41.

Sacks, Elisha P. 1991. Automatic analysis of one-parameter planar odes by intelligent numerical simulation. *Artificial Intelligence* 48.

Yip, Kenneth 1991. *KAM: A System for Intelligently Guiding Numerical Experimentation by Computer.* Artificial Intelligence Series. MIT Press.

Zhao, Feng 1991. Extracting and representing qualitative behaviors of complex systems in phase spaces. In *Proceedings IJCAI-91.* International Joint Conference on Artificial Intelligence.

Zwillinger, Daniel 1992. *Handbook of differential equations.* Academic Press., 2nd edition. ISBN 0-12-784391-4.