

Optimizing Talbot's Contours for the Inversion of the Laplace Transform

J.A.C. Weideman
Department of Applied Mathematics
University of Stellenbosch
South Africa

Talbot's method for the numerical inversion of the Laplace Transform consists of numerically integrating the Bromwich integral on a special contour by means of the trapezoidal or midpoint rules. In this paper we address the issue of how to choose the parameters that define the contour, for the particular situation when parabolic PDEs are solved. In the process the well known subgeometric convergence rate $O(e^{-c\sqrt{N}})$ of this method is improved to the geometric rate $O(e^{-cN})$, with N the number of nodes in the integration rule. The value of the maximum decay rate c is explicitly determined. Numerical results involving two versions of the heat equation are presented. With the choice of parameters derived here, the rule-of-thumb is that to achieve an accuracy of $10^{-\ell}$ at any given time t , the associated elliptic problem has to be solved no more than ℓ times.

Key words and phrases: Laplace Transform, Talbot's Method, Trapezoidal Rule

[Supported by the National Research Foundation in South Africa under grant NRF5289.]

Oxford University Computing Laboratory
Numerical Analysis Group
Wolfson Building
Parks Road
Oxford, England OX1 3QD
E-mail: andwn@comlab.oxford.ac.uk

March, 2005

Contents

1	Introduction	3
2	Analysis of the Scalar Problem	7
3	New Parameters for the Contour	11
4	Computing the Optimal Parameters	16
5	Application to PDEs	20
6	Comparisons	27

1 Introduction

The Laplace Transform is a classical technique for solving linear differential equations. For computational work, however, this approach never really became popular as for many years numerical analysts tended to focus on discretization methods such as finite differences and finite elements, possibly combined with linear multistep or Runge-Kutta formulas for integration in time. We conjecture that this lack of interest shown by numerical analysts in the Laplace Transform is partly due to the following two factors.

First, the Laplace Transform restricts one to linear differential equations and in many applications one ultimately wishes to solve nonlinear problems. Second, the Laplace Transform and particularly its numerical inversion has a reputation for being a computational challenge. This has to do with the fact that the inverse problem is by nature ill-conditioned when the transform is known only as a real-valued function [24]. When the transform can be sampled in the complex plane the conditioning seems better, but then complex-valued arithmetic is of course necessary.

Despite these apparent drawbacks of the Laplace Transform, there has been a recent resurgence of the technique, particularly in the area of linear parabolic PDEs; examples are the papers [5, 10, 14, 15, 18, 19], all of which appeared in the year 2000 or later. Why there has been this sudden interest one can only speculate, but the fact that MATLAB and other modern computational environments make complex arithmetic as easy to work with as real arithmetic may have something to do with it.

To introduce the problem, consider the linear system of ordinary differential equations

$$\frac{d\mathbf{f}}{dt} = A\mathbf{f}, \quad \mathbf{f}(0) = \mathbf{f}_0, \quad (1.1)$$

where A is an $M \times M$ real matrix, $\mathbf{f}(t)$ an $M \times 1$ real vector, and \mathbf{f}_0 the initial condition. We are primarily interested in the case where A is the result of semi-discretization (i.e., discretization in space) of a parabolic PDE (examples are given in §4). We assume, therefore, that the eigenvalues of A are real and negative.

The formal solution to (1.1) is

$$\mathbf{f}(t) = \exp(At) \mathbf{f}_0,$$

and this reduces the problem to that of computing the matrix exponential of a (typically) large matrix. To be more precise, we need to compute the product of the matrix exponential and a vector, which can be done without actually computing the matrix exponential itself [16].

The authors of [5, 10, 14, 15, 18, 19] all compute this product by numerically approximating the inverse Laplace Transform

$$\mathbf{f}(t) = \frac{1}{2\pi i} \int_{\Gamma} e^{zt} \mathbf{F}(z) dz, \quad (1.2)$$

where

$$\mathbf{F}(z) = (zI - A)^{-1} \mathbf{f}_0.$$

In this formula, known as the Bromwich integral, I is the $M \times M$ identity matrix and Γ is the contour of integration. At least initially, Γ is the Bromwich line $\operatorname{Re} z = \sigma$, where the parameter σ should be large enough that all eigenvalues of A lie in the half-plane $\operatorname{Re} z < \sigma$.

The typical approach is to deform the Bromwich line into a curve that begins and ends in the left-half plane, such that $\operatorname{Re} z \rightarrow -\infty$ on the contour; see for example Figures 1 and 10 below. Owing to the exponential factor e^{zt} , the integrand decays rapidly on such a contour, and if the contour is smooth this turns the problem into one of the classic situations where the trapezoidal rule converges extraordinarily rapidly [6, 12, 23, 27].

The articles [5, 10, 14, 15, 18, 19] differ with respect to the choice of the integration contour Γ , and how this contour is parameterized before approximating the integral with the trapezoidal rule. We present a summary of various contours and their associated convergence rates in §6.

None of the above references seriously considers Talbot's method [20], rated in some circles as one of the best methods for inverting the Laplace Transform; see [7]. (The method is mentioned in [14] and [19], but is neither implemented nor analyzed there.) For reasons discussed in §6, Talbot's method may not be suitable when part of the spectrum of A moves off the negative real axis. But for pure parabolic problems the method is very accurate as the numerical results of this paper will testify.

Talbot's method is based on a contour which, after parameterization, can be expressed as

$$\Gamma : \quad z(\theta) = \sigma + \mu (\theta \cot \theta + \nu i \theta), \quad -\pi \leq \theta \leq \pi. \quad (1.3)$$

Here σ , $\mu (> 0)$ and $\nu (> 0)$ are real parameters that determine the geometry of the curve. At the very least, for the eigenvalues of A to be enclosed by the contour one needs $z(0) > \lambda$, where λ is the largest eigenvalue of A i.e.,

$$\sigma + \mu > \lambda. \quad (1.4)$$

A typical contour is shown in Figure 1.

A closely related contour is obtained by replacing the function $\theta \cot \theta$ in (1.3) with the first two terms in its partial fraction expansion,

$$\Gamma : \quad z(\theta) = \sigma + \mu \left(1 + \frac{2\theta^2}{\theta^2 - \pi^2} + \nu i \theta \right), \quad -\pi \leq \theta \leq \pi. \quad (1.5)$$

This contour is equivalent to one mentioned in Talbot's original paper [20], from which we quote "...and indeed such functions can give good results, though their potentialities have not yet been explored."

It will turn out that the contour (1.5) is easier to analyze than (1.3), so for much of the paper we shall focus on the second contour. We shall also show, however, that the first Talbot contour is superior in terms of accuracy.

Using either (1.3) or (1.5), the Bromwich integral (1.2) can be expressed as

$$\mathbf{f}(t) = \frac{1}{2\pi i} \int_{-\pi}^{\pi} e^{z(\theta)t} \mathbf{F}(z(\theta)) z'(\theta) d\theta, \quad (1.6)$$

where, respectively,

$$z'(\theta) = \mu (\cot \theta - \theta \csc^2 \theta + \nu i) \quad \text{or} \quad z'(\theta) = \mu \left(-\frac{4\pi^2 \theta}{(\theta^2 - \pi^2)^2} + \nu i \right).$$

The integral (1.6) is typically approximated by the trapezoidal rule on a uniform partition of $[-\pi, \pi]$. Instead, we prefer to use the equally accurate midpoint rule with an even number of intervals, say $2N$. This is a practical choice that avoids sampling the integrand at the removable singularity at $\theta = 0$, as well as at the essential singularities at $\theta = \pm\pi$.

We hence define the grid

$$\theta_k = (2k + 1) \frac{\pi}{2N}, \quad k = -N, \dots, N - 1, \quad (1.7)$$

and denote the approximation to (1.6) by

$$\mathbf{f}_N(t) = \frac{1}{2Ni} \sum_{k=-N}^{N-1} e^{z(\theta_k)t} z'(\theta_k) \mathbf{F}_k,$$

or

$$\mathbf{f}_N(t) = \frac{1}{N} \operatorname{Im} \left\{ \sum_{k=0}^{N-1} e^{z(\theta_k)t} z'(\theta_k) \mathbf{F}_k \right\}, \quad (1.8)$$

if symmetry is used. Here the vectors $\mathbf{F}_k \equiv \mathbf{F}(z(\theta_k))$ are solved from

$$(z(\theta_k) I - A) \mathbf{F}_k = \mathbf{f}_0, \quad k = 0, \dots, N - 1. \quad (1.9)$$

Unless A is sparse, the solution of the N linear systems (1.9) represents the bulk of the computational cost of the algorithm. It should however be noted that the systems (1.9) can be solved independently and in parallel [12, 18]. In addition, it is possible to solve all N systems (1.9) efficiently using a single Hessenberg or Schur decomposition of A ; see Problem P7.4.2 in [11, p. 350].

In this paper, we shall aim to optimize the convergence rate $\mathbf{f}_N(t) \rightarrow \mathbf{f}(t)$ as $N \rightarrow \infty$, by selecting the parameters (σ, μ, ν) in (1.3) to be asymptotically optimal. This is achieved by making σ and μ both proportional to the ratio N/t . By doing so, a geometric convergence rate, $O(e^{-cN})$ as $N \rightarrow \infty$, can be obtained. It is well known that Talbot's method with fixed (and therefore suboptimal) parameters converges at a subgeometric rate of $O(e^{-c\sqrt{N}})$ [20]. The best convergence rate reported to date is still subgeometric, namely the $O(e^{-cN/\log N})$ achieved on the hyperbolic contour of [14]. For a summary of convergence rates reported in the literature, see Table 2.

To conclude this introduction, and to illuminate the role of the parameters (σ, μ, ν) , we present Figure 1. The contour depicted is (1.3), with an arbitrary choice of parameters (σ, μ, ν) . As $\operatorname{Re} z \rightarrow -\infty$, the contour approaches the horizontal asymptotes $\operatorname{Im} z = \pm\pi\mu\nu$, as indicated by the dash-dot lines. The dots represent the images of the midpoint abscissas (1.7) in the z -plane ($N = 16$, i.e., 32 nodes in total). Note that the nodes have

constant spacing in the imaginary direction. In fact, this constant is $\pi\mu\nu/N$ which is therefore roughly the spacing of the nodes near the real axis. Finally, the extreme two nodes (i.e., the two of largest modulus) are $z(\theta_{-N})$, given asymptotically by

$$z(\theta_{-N}) = \sigma + \mu - 2\mu N - \pi\mu\nu i + O(N^{-1}), \quad N \rightarrow \infty,$$

and its conjugate, $z(\theta_{N-1})$.

The role of the parameters (σ, μ, ν) should now be clear. The first, σ , is just a translation that moves the contour to the left or right. The extreme two nodes, $z(\theta_{-N})$ and $z(\theta_{N-1})$, are controlled by μ : the larger this parameter is, the farther into the left half-plane these outlying nodes move. The width of the contour, or equivalently, the relative spacing of nodes near the real axis, are both proportional to $\mu\nu$.

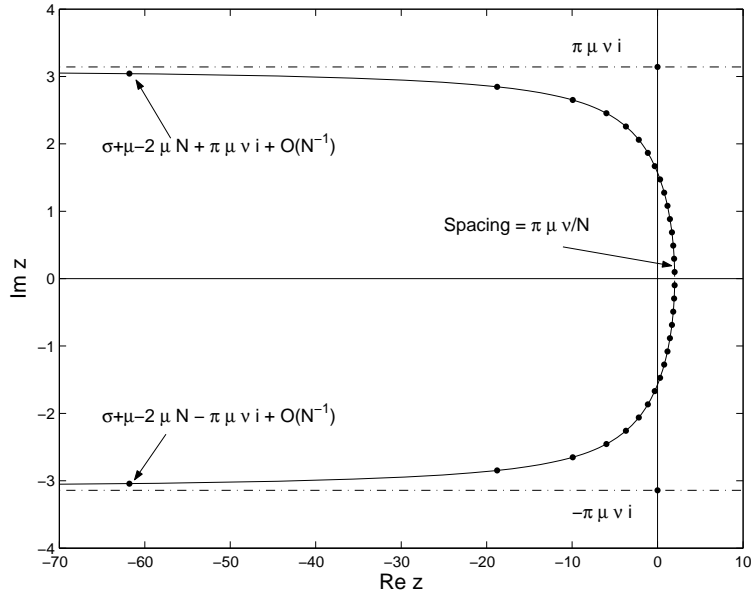


Figure 1: Talbot's contour (1.3) with parameter values $\sigma = 0$, $\mu = 2$, $\nu = 0.5$. The dots are the images in the z -plane of the midpoint abscissas (1.7).

The outline of the paper is as follows. In §2, we rederive the well-known $O(e^{-c\sqrt{N}})$ convergence rate of Talbot's method, by analyzing the scalar model problem

$$f(t) = e^{\lambda t}, \quad F(z) = \frac{1}{z - \lambda}. \quad (1.10)$$

This analysis suggests the re-parameterization of the Talbot contour that we alluded to above, namely to make both σ and μ proportional to N/t . A saddle point method is used in §3 to demonstrate that this rescaling of parameters, if done judiciously, leads to the improved convergence rate $O(e^{-cN})$. In §4, we determine the value of ν and the proportionality constants in $\sigma \propto N/t$, $\mu \propto N/t$, that will maximize the decay rate, c . We hence obtain the attractive convergence rates $O(e^{-1.90N})$ and $O(e^{-1.73N})$, respectively, for

the two versions of the Talbot contour (1.3) and (1.5). A further improvement, which involves omitting those outlying nodes on the Talbot contour that make a negligible contribution to the midpoint sum, improves these two convergence rates to effectively $O(e^{-2.41N})$ and $O(e^{-2.53N})$. The theory of §3–§4 is tested on two parabolic PDEs in §5. In §6, we discuss a few contours that can be used as alternative to Talbot's, and we also contrast the parameter suggestions of this paper with those made by Talbot in the original paper [20].

2 Analysis of the Scalar Problem

We suppose the matrix A , which may or may not be symmetric, has real and negative eigenvalues, λ_j , corresponding to a complete set of eigenvectors, \mathbf{v}_j , $j = 1, \dots, M$. If one expands the initial condition as a linear combination of eigenvectors,

$$\mathbf{f}_0 = c_1 \mathbf{v}_1 + \dots + c_M \mathbf{v}_M,$$

then it is a routine exercise to show that (1.2) can be expressed as

$$\mathbf{f}(t) = \frac{c_1}{2\pi i} \int_{\Gamma} \frac{e^{zt}}{z - \lambda_1} dz + \dots + \frac{c_M}{2\pi i} \int_{\Gamma} \frac{e^{zt}}{z - \lambda_M} dz.$$

Applying Talbot's method to the right-hand side is therefore equivalent to applying it to the scalar problem (1.10), where λ represents a real and negative eigenvalue of A . For any fixed t , we shall restrict attention to λ in the range

$$|\lambda t| = O(1), \quad M \rightarrow \infty.$$

We consider this sufficient, because in the actual solution,

$$\mathbf{f}(t) = c_1 e^{\lambda_1 t} \mathbf{v}_1 + \dots + c_M e^{\lambda_M t} \mathbf{v}_M,$$

modes that satisfy $|\lambda t| \gg 1$ are negligible.

Our task is therefore to estimate the error when approximating the integral

$$f(t) = \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{e^{z(\theta)t}}{z(\theta) - \lambda} z'(\theta) d\theta \quad (2.1)$$

with the midpoint rule, with $z(\theta)$ defined by (1.3) or (1.5), and $\lambda < 0$. To start, we recall an error formula for the midpoint rule.

Consider an integral on $[-\pi, \pi]$ and its midpoint rule approximation

$$I(g) = \int_{-\pi}^{\pi} g(\theta) d\theta, \quad M_N(g) = \frac{\pi}{N} \sum_{k=-N}^{N-1} g(\theta_k), \quad (2.2)$$

where the nodes θ_k are defined by (1.7). (Note that we continue to use $2N$ intervals, i.e., step-size $h = \pi/N$.) Suppose that the function $g(\theta)$ has an absolutely convergent Fourier series

$$g(\theta) = \sum_{k=-\infty}^{\infty} c_k e^{ik\theta},$$

with coefficients given by

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} g(\theta) e^{-ik\theta} d\theta, \quad k = 0, \pm 1, \pm 2, \dots$$

Then it is possible to insert these formulas into (2.2), followed by termwise integration and summation, to obtain

$$I(g) = 2\pi c_0, \quad M_N(g) = 2\pi c_0 + 2\pi \sum_{\substack{\ell=-\infty \\ \ell \neq 0}}^{\infty} (-1)^\ell c_{2\ell N}.$$

The error is therefore given by

$$I(g) - M_N(g) = -2\pi \sum_{\substack{\ell=-\infty \\ \ell \neq 0}}^{\infty} (-1)^\ell c_{2\ell N}. \quad (2.3)$$

(The trapezoidal rule error would be similar, except for the absence of the $(-1)^\ell$ factor; see [27].)

When the periodic extension of $g(\theta)$ is infinitely differentiable on $[-\pi, \pi]$, the Fourier coefficients c_k decay rapidly. In fact, repeated integration by parts can then be used in the usual manner to establish $c_k = O(|k|^{-m})$ for each positive integer m . In such cases a good estimate for the error can usually be obtained by retaining only the leading two terms in the error formula (2.3):

$$\begin{aligned} I(g) - M_N(g) &\sim 2\pi (c_{-2N} + c_{2N}) \\ &= \int_{-\pi}^{\pi} g(\theta) e^{-2Ni\theta} d\theta + \int_{-\pi}^{\pi} g(\theta) e^{+2Ni\theta} d\theta. \end{aligned}$$

This differs from the trapezoidal error estimate only by a sign.

One may apply the above estimate to the special integral (2.1). The factor $e^{z(\theta)t}$ decays rapidly as $\theta \rightarrow \pm\pi$, which ensures infinite differentiability of the 2π -periodic extension of the integrand. We therefore propose to analyze the error estimate,

$$f(t) - f_N(t) \sim E_N^-(t) + E_N^+(t), \quad N \rightarrow \infty,$$

where

$$E_N^\pm(t) = \frac{1}{2\pi i} \int_{-\pi}^{\pi} \frac{e^{z(\theta)t \pm 2iN\theta}}{z(\theta) - \lambda} z'(\theta) d\theta,$$

or

$$E_N^\pm(t) = \frac{1}{\pi} \text{Im} \left\{ \int_{-\pi}^0 \frac{e^{z(\theta)t \pm 2iN\theta}}{z(\theta) - \lambda} z'(\theta) d\theta \right\} \quad (2.4)$$

if symmetry is used. We shall keep both $t > 0$ and $\lambda < 0$ fixed, as well as the parameters σ , μ and ν in the contours (1.3) or (1.5); our interest is the behaviour of these two error integrals as $N \rightarrow \infty$.

We digress for a moment to point out that error estimates such as (2.4) were used to good effect by F. Lin, to numerically predict optimal parameters for Talbot's contour [13]. A wide range of transforms was considered there, not just the $F(z) = 1/(z - \lambda)$ considered here.

Rather than using numerical optimization, as was done in [13], we shall instead use analytic methods, and in particular the saddle point method, to estimate the two integrals (2.4). Since this analysis is primarily used to justify the form of the rescaling of parameters in §3, and not in the determination of the actual optimal numbers itself, we keep the details to a minimum. (For general discussions of the saddle point method, we refer to [1, Sect. 6.4] and [3, Sect. 6.6], and for specific applications to problems similar to the one we are studying here, we refer to [4, 8, 25, 26].)

We define the two functions

$$h_\pm(\theta) = z(\theta)t \pm 2iN\theta.$$

The saddle points are defined as the solutions to $h'_\pm(\theta) = 0$, two equations that can be solved by asymptotic iteration. Expecting saddle points near the essential singularities, we use $\theta = -\pi$ as starting value in each case. We thus obtain two movable saddle points, θ_+ and θ_- , respectively located in the upper and lower θ -planes. To leading order

$$\theta_\pm = -\pi + \sqrt{\frac{\pi\mu t}{2N}} e^{\pm\pi i/4} + O(N^{-3/2}), \quad N \rightarrow \infty,$$

an estimate valid for both contours (1.3) and (1.5).

Assuming the steepest descents analysis can be completed successfully, there will be saddle point contributions to the integral (2.4) of the form

$$E_N^+(t) = O\left(e^{\text{Re}\{h_+(\theta_+)\}}\right), \quad E_N^-(t) = O\left(e^{\text{Re}\{h_-(\theta_-)\}}\right).$$

We therefore calculate $h_\pm(\theta_\pm)$ and obtain, as $N \rightarrow \infty$,

$$\text{Re}\{h_\pm(\theta_\pm)\} = (\sigma + \mu)t - 2\sqrt{\pi\mu tN} + O(N^{-1/2})$$

in the case of contour (1.3), and

$$\text{Re}\{h_\pm(\theta_\pm)\} = \left(\sigma + \frac{5}{2}\mu\right)t - 2\sqrt{\pi\mu tN} + O(N^{-1/2})$$

in the case of contour (1.5). We note that there will also be asymptotic contributions from the singularities associated with the roots of $z(\theta) = \lambda$, but these are of the form $O(e^{-cN})$ and therefore subdominant to the saddle point contributions.

These informal arguments predict, therefore, that the essential convergence rate is of the form

$$E_N(t) = O\left(e^{(\sigma+\mu)t-2\sqrt{\pi\mu tN}}\right), \quad N \rightarrow \infty, \quad (2.5)$$

in the case of contour (1.3), and

$$E_N(t) = O\left(e^{(\sigma+\frac{5}{2}\mu)t-2\sqrt{\pi\mu tN}}\right), \quad N \rightarrow \infty, \quad (2.6)$$

in the case of contour (1.5).

Before proceeding to a numerical check of (2.5) and (2.6), we make two remarks. First, similar results were obtained by Talbot [20, eq. (15)], who proved that

$$E_N(t) = O\left(N^2 e^{(\sigma+a\mu)t-b\sqrt{N}}\right)$$

using a method different from ours. The constants a and b were not determined explicitly, however. Second, the factor $5/2$ that appears in (2.6) suggests that the error constant is larger for the contour (1.5) than for (1.3). This indicates that (1.3) is superior to (1.5), a conclusion supported by the following numerical experiment.

In Figure 2, we show typical error curves computed with Talbot's method. Note the logarithmic and square root scales on the vertical and horizontal axes, respectively, chosen in order that the error estimates (2.5)–(2.6) should appear as straight lines. As for the scalloped shape of the error curves, we remark that the formulas underlying (2.5) and (2.6) both contain multiplicative factors of the form $\cos(2\sqrt{\pi\mu tN} + \beta)$ that are suppressed by the order notation. This accounts for the oscillations in the error curves, with half-period $\frac{1}{2}\sqrt{\pi/\mu t}$, a value of about 0.4 for the parameters used here and therefore consistent with the figure. Finally, according to these asymptotic error estimates the error when using contour (1.5) should be larger by a factor $e^{3\mu t/2}$, which is about 1.8×10^3 here. This is in agreement with the three orders of magnitude gap between the upper and lower sets of curves.

The arguments presented above, as well as the asymptotic analysis of the next section, indicate that as $N \rightarrow \infty$, the errors become independent of λ . (In fact, in §3 we shall effectively set $\lambda = 0$.) Of course this does not mean that the errors are independent of λ for finite values of N , and in the end it will be important to carry out numerical experiments to confirm that such an approximation was justified. Indeed, when one plots error curves such as those in Figure 2 using a larger value of $|\lambda|$, say $\lambda = -5$, the $O(e^{-cN})$ contribution dominates the saddle point contributions $O(e^{-c\sqrt{N}})$ until N is quite large.

The reader familiar with MATLAB will recognize the use of that software package's plotting routines in Figure 2. He or she might also be perplexed by the observation of accuracies as small as 10^{-30} , far smaller than MATLAB's machine epsilon of about 2.2×10^{-16} . Here, and elsewhere in the paper where multiprecision arithmetic was needed, computations were done in Maple with the results exported to MATLAB for plotting.

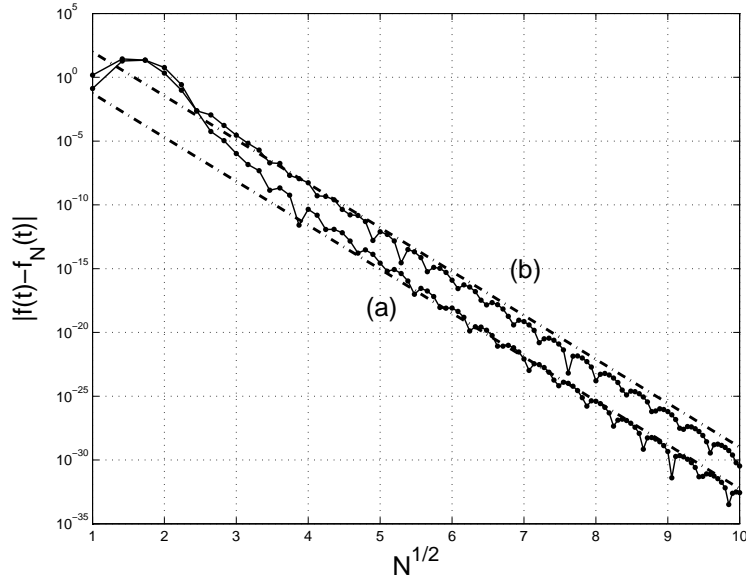


Figure 2: Absolute errors when Talbot's method with fixed parameters is applied to the model problem (1.10), using (a) the contour (1.3), and (b) the contour (1.5). The thinner, dotted curves are the numerically computed errors, and the thicker dash-dot lines represent the theoretical error estimates (2.5) and (2.6). The parameter values are $\sigma = 0$, $\mu = 5$, $\nu = 1$, $t = 1$, $\lambda = -1$. Note the $N^{1/2}$ horizontal axis, which confirms the $O(e^{-c\sqrt{N}})$ convergence rate (2.6). This will be improved to $O(e^{-cN})$ in the next section by choosing optimal parameters.

3 New Parameters for the Contour

The error estimates (2.5)–(2.6) suggest a way of improving the convergence rate of Talbot's method. By rescaling the parameters to

$$\sigma = -s\frac{N}{t}, \quad \mu = m\frac{N}{t}, \quad \nu = n, \quad (3.1)$$

it should be possible to improve the subgeometric convergence rate, $e^{-c\sqrt{N}}$, to pure geometric convergence, e^{-cN} , provided the parameters (s, m, n) are judiciously selected. Here $m > 0$, $n > 0$, and in accordance with (1.4) we shall require that

$$s < m - \frac{\lambda t}{N}. \quad (3.2)$$

The constant λ is defined in (1.10), which we continue to use as model problem.

Because the parameters become dependent on t , so does the contour and hence also the integration nodes. This means that the N linear systems (1.9) need to be solved for each value of t , which is inefficient. We therefore intend this rescaling to be used when the solution is required at only a few values of t . In §6, however, we shall explore the

idea of fixing t at a particular value, and then using the corresponding parameters (3.1) to compute the solution at nearby values of t , thereby improving efficiency greatly.

Using the new parameters (3.1) we define $\zeta(\theta) = (t/N) z(\theta)$, i.e.,

$$\zeta(\theta) = -s + m \left(\theta \cot \theta + i n \theta \right), \quad (3.3)$$

in the case of the contour (1.3), and

$$\zeta(\theta) = -s + m \left(1 + \frac{2\theta^2}{\theta^2 - \pi^2} + i n \theta \right), \quad (3.4)$$

in the case of the contour (1.5). The error integral (2.4) therefore becomes

$$E_N^\pm(t) = \frac{1}{\pi} \text{Im} \left\{ \int_{-\pi}^0 \frac{e^{N g_\pm(\theta)}}{\zeta(\theta) - \lambda t/N} \zeta'(\theta) d\theta \right\}, \quad (3.5)$$

where

$$g_\pm(\theta) = \zeta(\theta) \pm 2i\theta.$$

We apply the saddle point method to the integral (3.5). The idea is to deform the interval of integration, $[-\pi, 0]$, to a special contour in the complex θ -plane on which the integral can be estimated accurately. By Cauchy's theorem such a deformed contour will be permissible as long as it starts at $\theta = -\pi$, terminates at $\theta = 0$, and does not cross any singularities of the integrand in between.

Suitable contours are steepest descent curves, defined by $\text{Im}\{g_\pm(\theta)\} = \text{constant}$, for these remove the oscillations from the integrands in (3.5). The constant is chosen to pass through the saddle points, since the integrand then exhibits a sharp peak which allows for an application of Laplace's method to estimate the integral [1, Sect. 6.4], [3, Sect. 6.6]. The saddle points, $\theta = \theta_+$ and $\theta = \theta_-$, are respectively defined as solutions to

$$g'_+(\theta) = 0, \quad g'_-(\theta) = 0. \quad (3.6)$$

To ensure analyticity of the integrand, one needs to take into consideration the singularities associated with the vanishing of the denominator in (3.5), i.e., the zeros of $\zeta(\theta) = \lambda t/N$. In view of the discussion in the first paragraph of §2, we shall assume $|\lambda t| \ll N$ and ignore the right-hand side of this equation. We therefore define the critical points, $\theta = \theta_*$, as the roots of

$$\zeta(\theta) = 0. \quad (3.7)$$

This is the same as setting $\lambda = 0$, and in accordance with (3.2) we shall therefore only consider $m > s$.

To apply the saddle point method, we need an idea of where the critical points and saddle points are. We restrict ourselves to the modified contour (1.5), i.e., we assume $\zeta(\theta)$ is defined by (3.4). For this contour the equations (3.6) and (3.7) reduce to polynomial equations of degrees 4 and 3, respectively, which can be analyzed without great difficulty. In principle the methods of Cardano and Vieta can be applied to obtain explicit solutions,

but these results proved to be too complicated to extract useful information from, and we opted for a more elementary approach.

In order to work with equations with real coefficients, we introduce the variable ϕ by $\theta = i\phi$. After denominators have been cleared, (3.6) can be factored into

$$(mn \pm 2)(\phi^2 + \pi^2)^2 = 4m\pi^2\phi, \quad (3.8)$$

and (3.7) into

$$(mn\phi + s - m)(\phi^2 + \pi^2) = 2m\phi^2. \quad (3.9)$$

Assuming $m > s \geq 0$, and working with these two representations we were able to establish the following properties of the roots of (3.6)–(3.7).

Starting with the cubic equation (3.9), it is readily established that it always has a positive real root. The remaining two roots may either be real as well, or occur as a conjugate pair. In the latter case the real part of these roots is positive, and the imaginary part is bounded in absolute value by π . Transplanting this information from the ϕ variable to θ , we deduce that the three critical points θ_* defined by (3.7) are all in the upper half-plane, with real parts in the interval $(-\pi, \pi)$. At least one root lies on the positive imaginary axis. The other two roots may be pure imaginary as well, or they may be located symmetrically with respect to the imaginary axis. (In the next section, we shall conjecture that the optimal configuration occurs when this pair of roots bifurcates from a double root on the imaginary axis into the complex plane.)

Turning to (3.8), one notices that it is a quartic equation with real coefficients that is missing its cubic term. The typical configuration of roots is therefore one in each quadrant of the complex θ -plane, at equal distances from the real axis. The exception is when (3.8) admits real roots, i.e., saddle points on the imaginary θ -axis, but we shall ignore this case as it corresponds to suboptimal choices of the parameters (s, m, n) as will be demonstrated below.

Figure 3 shows a typical configuration of critical and saddle points. The roots of (3.6) are represented by +’s and ×’s (corresponding to the + and − signs, respectively), and the roots of (3.7) are plotted as the *’s.

We propose a saddle point analysis, based on the contours shown in the figure. The Γ_{\pm} are the curves of steepest descent $\text{Im}\{g_{\pm}(\theta)\} = \text{constant}$. Writing $\theta = x + yi$, they can be expressed as

$$\Gamma_{\pm} : \quad \frac{4m\pi^2 xy}{(x^2 - y^2 - \pi^2)^2 + 4x^2 y^2} - (mn \pm 2)x = c_{\pm}.$$

The constant c_{\pm} is determined by the requirement that each Γ_{\pm} passes through its corresponding saddle point, θ_+ or θ_- , as defined by (3.6).

In the lower half-plane, Γ_- starts at $\theta = -\pi$, passes through θ_- , and continues to $\theta = -i\infty$. This is valid since the integrand in (3.5), with minus sign, approaches zero as $\theta \rightarrow -i\infty$. The contour is then closed at $-i\infty$ and returns to the origin via the negative imaginary θ -axis. On this axis the contribution can be ignored, since the integrand is real. The error $E_N^-(t)$ is therefore solely determined by the saddle point contribution,

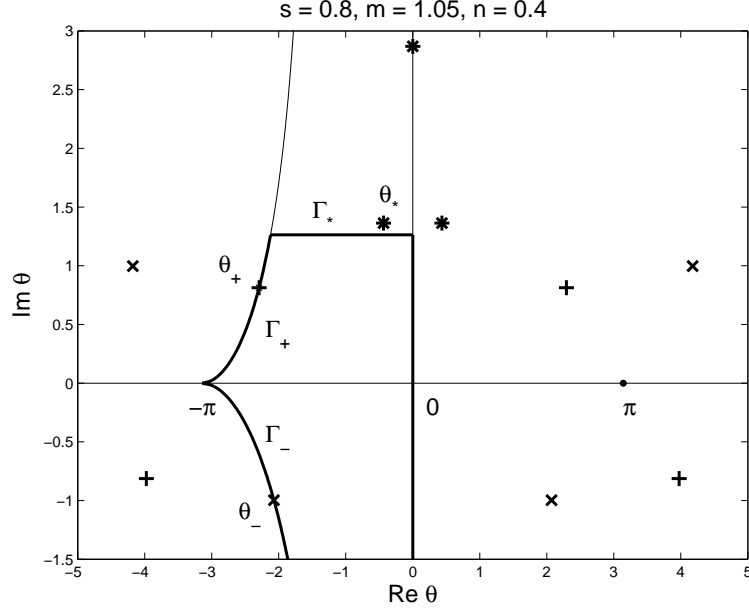


Figure 3: Saddle points, θ_- and θ_+ , critical points θ_* , and steepest descent contours used in deriving the error estimates (3.10), (3.11), (3.12).

which can be computed in the usual manner as [1, Sect. 6.4], [3, Sect. 6.6]

$$E_N^-(t) = O(e^{d_- N}), \quad d_- = \operatorname{Re}\{g_-(\theta_-)\}. \quad (3.10)$$

In the upper half-plane a similar approach is used, except for the fact that the critical points θ_* have to be taken into account. The contour Γ_+ is not continued to $\theta = +i\infty$, as it will not be possible to return to the origin without crossing the singular points $\theta = \theta_*$. To ensure that the integrand remains analytic, we introduce a third contour, Γ_* , that branches off from Γ_+ and has constant imaginary part, say $\operatorname{Im}\{\Gamma_*\} = b$. Typically, the value of b would be determined by the critical point θ_* nearest to the real axis. By letting Γ_* approach such a limiting θ_* from below, it is possible to establish

$$E_N^+(t) = O(e^{d_* N}), \quad d_* = \operatorname{Re}\{g_+(\theta_*)\}. \quad (3.11)$$

If θ_+ lies below Γ_* , a saddle point contribution similar to (3.10),

$$E_N^+(t) = O(e^{d_+ N}), \quad d_+ = \operatorname{Re}\{g_+(\theta_+)\}, \quad (3.12)$$

is to be added to (3.11).

In our numerical experiments, the total error was dominated either by (3.10) or by (3.11). We have not found a set of parameters (s, m, n) for which (3.12) dominates, but neither have we tried to prove that this is impossible.

In Figure 4 below, we offer numerical verification of these error estimates. In the Appendix, the corresponding values of saddle points, critical points and expected convergence rates are summarized. We have picked sets of parameter values (s, m, n) for

which (a) the saddle point contribution (3.10) dominates, (b) the critical point contribution (3.11) dominates, and (c) where these two contributions are approximately equal. (In the next section we shall conjecture that this is the optimal situation.) Also shown, as the dash-dot curves are the predicted convergence rates, i.e., the maximum of (3.10) and (3.11). Here we should point out that these estimates are, of course, asymptotic and much information is suppressed by the order notation of (3.10), (3.11), and (3.12). Therefore, in some cases N has to be large for the estimate to become valid. This can be seen in part (b) of Figure 4, for example, where N needs to be greater than 70, roughly, before (3.11) becomes evident.

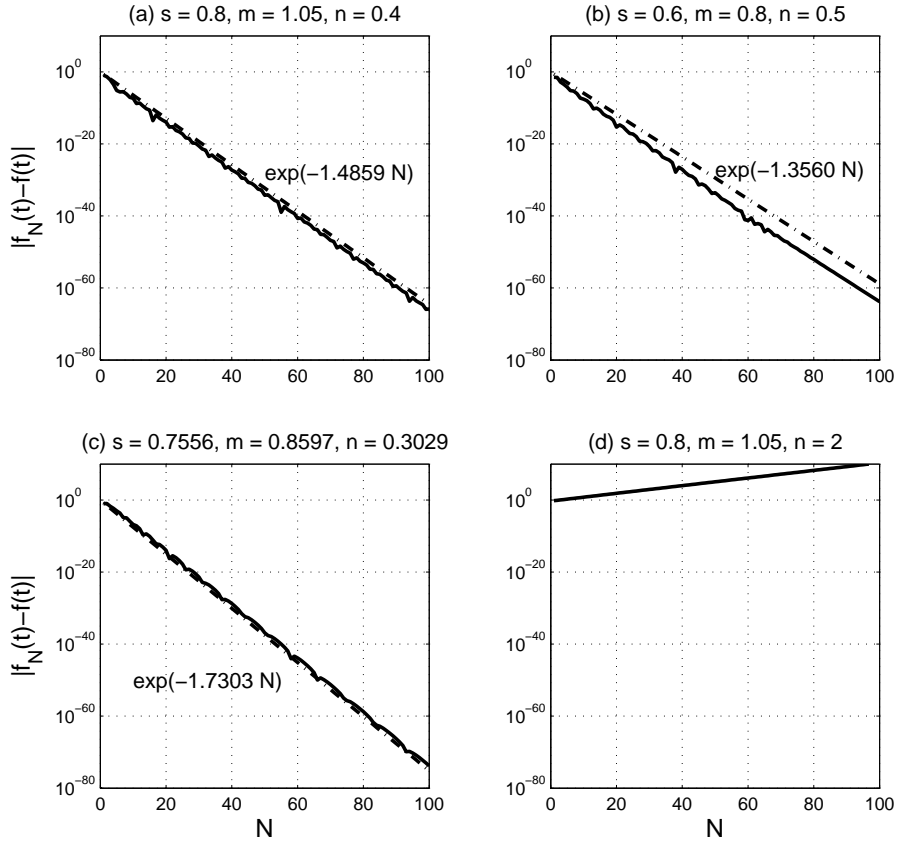


Figure 4: Convergence curves for Talbot's method applied to the model problem (1.10), with $\lambda = -1$, $t = 1$, and using contour (1.5). The solid curves represent the actual errors, and the dash-dot curves the theoretical estimates as summarized in the table in the Appendix.

The purpose of Case (d) in Figure 4 is to demonstrate that for certain choices of parameters (s, m, n) the method actually diverges. With this choice the saddle point θ_- does not lie in the correct (third) quadrant and the saddle point analysis cannot be completed as described above. Rather than trying to determine the allowable set of parameters for which the method is guaranteed to converge, we shall rather focus our

attention on finding the choice that gives the quickest convergence rate, which is the topic of the next section.

4 Computing the Optimal Parameters

A first attempt at finding optimal parameters (s, m, n) , was based on a numerical optimization strategy, involving the objective function

$$F(s, m, n) \equiv \max \{d_+, d_-, d_*\} = \text{minimum}. \quad (4.1)$$

Here d_+, d_-, d_* are the decay constants in the error estimates (3.12), (3.10), and (3.11), respectively. For each set of parameters (s, m, n) , the value of F can be computed by first solving (3.6) and (3.7) to obtain θ_+, θ_- , and θ_* . These values are then substituted into $g_{\pm}(\theta)$, to compute d_+, d_-, d_* as defined by (3.10)–(3.12).

In the case of contour (1.5), the equations (3.6) and (3.7) can be solved with polynomial rootfinding routines, and in this case MATLAB's function `roots` was used. In the case of contour (1.3) a complex Newton process was used. In solving (3.7), it requires some careful programming to select the correct root θ_* .

The problem (4.1) was solved using MATLAB's function `fminsearch`, a routine suitable for nonsmooth, unconstrained optimization. Aside from some mild ill-conditioning that will be explained below, this approach worked reasonably well.

In the case of contour (1.5), this approach yielded the parameters already presented as case (c) in Figure 4, namely

$$s = 0.7556, \quad m = 0.8597, \quad n = 0.3029. \quad (4.2)$$

The corresponding saddle and critical points are given by

$$\theta_+ = -2.3503 + 0.7620i, \quad \theta_- = -2.2315 - 0.8652i, \quad \theta_* = 0.8652i, \quad (4.3)$$

which correspond to

$$d_+ = -2.1524, \quad d_- = -1.7303, \quad d_* = -1.7303.$$

This predicts a convergence rate of

$$E_N(t) = O(e^{-1.7303N}), \quad N \rightarrow \infty. \quad (4.4)$$

Applying the same algorithm to the original Talbot contour (1.3), we obtained a significantly better convergence rate, namely

$$E_N(t) = O(e^{-1.8975N}), \quad N \rightarrow \infty. \quad (4.5)$$

This corresponds to parameter values

$$s = 0.4814, \quad m = 0.6443, \quad n = 0.5653, \quad (4.6)$$

with

$$\theta_+ = -2.5293 + 0.7435i, \quad \theta_- = -2.4158 - 0.9487i, \quad \theta_* = 0.9487i, \quad (4.7)$$

and

$$d_+ = -2.5048, \quad d_- = -1.8975, \quad d_* = -1.8975.$$

In Figure 5, we show the optimal configuration of the saddle and critical points associated with the original Talbot contour (1.3). The locations of θ_+ , θ_- and θ_* , are shown as the +, \times , and $*$ symbols in the top figure, respectively, and their images in the z -plane in the bottom figure. Also shown are the nodes used in the midpoint approximation, with $N = 16$ (i.e., 32 subintervals).

Examining the numerical results (4.6)–(4.7), we conjecture that in the optimal configuration,

- (a) θ_* is on the positive imaginary axis,
- (b) $\zeta'(\theta_*) = 0$,
- (c) $\text{Im}(\theta_*) = -\text{Im}(\theta_-)$, and
- (d) $d_+ < d_- = d_*$,

for both contours (1.3) and (1.5). All of these properties seem plausible, but we have not pursued rigorous proofs.

Property (b) indicates that θ_* is a double root of equation (3.7), which is the source of the ill-conditioning mentioned at the beginning of the section. Fortunately, assuming properties (a)–(d) to be true, the problem can be reformulated such that it becomes explicitly solvable. The details are as follows.

Using properties (a) and (c) above, we write

$$\theta_* = yi, \quad \theta_- = x - yi,$$

where $x < 0$ and $y > 0$. Because of property (d), we shall ignore θ_+ for now, and try to solve for x , y , s , m , and n from the following five (real) equations: the right-hand equality in eq. (3.6) (two real equations), eq. (3.7), property (b), and the equality in property (d).

Using a straightforward but tedious hand calculation the above 5×5 system was reduced to a 2×2 system involving only x and y . In the case of contour (1.5) this system is

$$\begin{aligned} x(P^2 - Q^2) + 2yPQ &= 0, \\ (5y^2 + \pi^2)(P^2 + Q^2) + P(y^2 + \pi^2)^2 &= 0, \end{aligned} \quad (4.8)$$

where

$$P = x^2 - y^2 - \pi^2, \quad Q = 2xy.$$

For further simplification we turned to Maple, which produced an explicit solution in the form

$$y = \pi\sqrt{v},$$

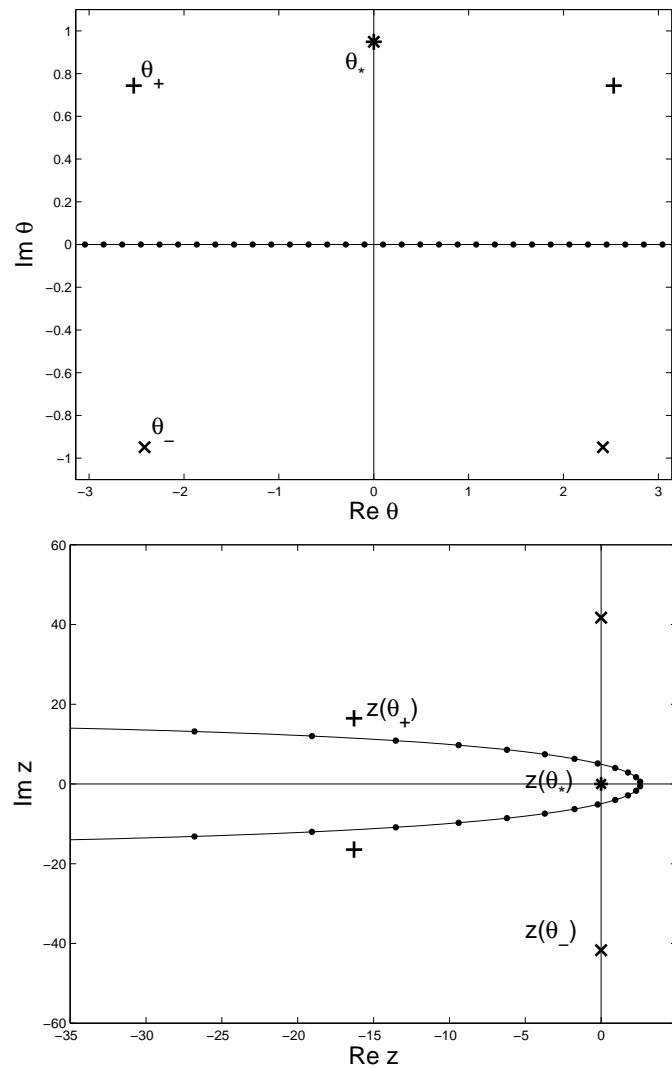


Figure 5: The optimal configuration of saddle points and critical points, shown as crosses, in the θ -plane (top) as well as their images in the z -plane (bottom). The dots are the nodes used in the midpoint rule, with $N = 16$. Note that in the bottom figure, four sets of nodes lie offscale (towards $\text{Re } z = -\infty$). The contour is given by (1.3), but the figure is qualitatively similar for (1.5).

where $v \approx 0.07584$ is the smallest positive root of the equation

$$41v^4 - 308v^3 - 98v^2 - 4v + 1 = 0.$$

With this value of v , the value of x may be computed as

$$x = -\frac{\pi}{4\sqrt{2}}\sqrt{41 - 209v - 1581v^2 + 205v^3}.$$

These formulas yield the values $y \approx 0.8652$ and $x \approx -2.2315$, as obtained numerically in (4.3). The values of s , m , and n follow from

$$n = \frac{4\pi^2 y}{(y^2 + \pi^2)^2}, \quad m = \frac{2(R^2 + S^2)}{4\pi^2(yR - xS) + n(R^2 + S^2)}, \quad s = m\left(\frac{3y^4 + \pi^4}{(y^2 + \pi^2)^2}\right),$$

where

$$R = P^2 - Q^2, \quad S = 2PQ.$$

In the case of contour (1.3) the analogue of the system (4.8) is

$$\begin{aligned} A - x(A^2 - B^2 + 1) - 2yAB &= 0 \\ xA + yB + y(\coth y - 2y \operatorname{csch}^2 y) &= 0, \end{aligned}$$

where we have defined $A = \operatorname{Re}\{\cot \theta_-\}$, $B = \operatorname{Im}\{\cot \theta_-\}$, i.e.,

$$A = \frac{\sin x \cos x}{\sin^2 x + \sinh^2 y}, \quad B = \frac{\sinh y \cosh y}{\sin^2 x + \sinh^2 y}.$$

A numerical solution of this system yields the value of $\theta_- = x - iy$ reported in (4.7). The values of the other parameters can be computed via

$$n = \coth y - y \operatorname{csch}^2 y, \quad m = \frac{2}{B + y(A^2 - B^2 + 1) - 2xAB + n}, \quad s = my^2 \operatorname{csch}^2 y.$$

In Table 1, we list the values, to sixteen significant digits, of the optimal parameters (s, m, n) as computed by these formulas. We emphasize, however, that the four place values already listed in (4.2) and (4.6) should be more than adequate in practice.

Table 1: Optimal parameter values $\sigma = -sN/t$, $\mu = mN/t$, and $\nu = n$ in the two Talbot contours.

	s	m	n
Contour (1.3)	0.4813750662515592	0.6442950965168778	0.5652899277355740
Contour (1.5)	0.7556411914150047	0.8597665100492216	0.3029436977358583

As verification that the parameters in Table 1 are indeed close to optimal, we offer Figure 6. There we show, as the thicker curves, the numerically computed errors $E_N(t)$

as functions of N , corresponding to parameters (4.2) and (4.6). Virtually on top of these curves and shown as dash-dot lines are the theoretical error estimates (4.4) and (4.5). To show the near-optimality of these curves, we have computed similar errors using a uniform sampling of parameter space $(s, m, n) \in (0, 1) \times (0, 1) \times (0, 1)$, with step-size 0.05 in each direction. (That is, $19^3 = 6859$ different parameter sets were used for each value of N .) The vertical “line segments” in the figure represent the range of these computed errors, with the minima and maxima indicated by the tiny horizontal bars.

We should not neglect to point out that if our sampling of parameter space were finer, some of the lower error bars in this figure could extend further down to 0. This will happen when the two error components, $E_N^+(t)$ and $E_N^-(t)$ in (2.4), are approximately of equal magnitude but of opposite sign. Such instances of fortuitous cancellation will however be rare when the matrix as opposed to the scalar problem is solved. We believe that Figure 6 represents solid evidence that the suggested parameter values in Table 1 are indeed asymptotically optimal.

To conclude this section, we point out a certain redundancy in the Talbot contour, as noted by L.N. Trefethen [22]. Recall Figure 5, where the optimal Talbot contour was shown for the case $N = 16$, and recall also that four pairs of nodes were located outside the frame of the figure, towards $\operatorname{Re} z = -\infty$. In fact, the contribution of each of these outlying nodes is negligible, as $|e^{z(\theta_k)t}| \leq e^{-1.8975N}$ when $|k| \geq 3N/4$. It therefore appears that virtually no accuracy is lost by including only the middle 75% of nodes, and discarding the outlying 25%. (A more careful calculation shows that the actual number of retained nodes should be about $0.7409N$.) With such a modification the efficiency is of course greatly enhanced. Quantitatively, the effective convergence rate jumps from $O(e^{-1.8975N})$ to about $O(e^{-1.8975N/0.75}) = O(e^{-2.53N})$.

In the case of the modified Talbot contour (1.5) the improvement is even more dramatic: One finds that only roughly 72% of the nodes need to be retained, which increases the effective convergence rate from $O(e^{-1.7303N})$ to about $O(e^{-2.41N})$.

It therefore appears that both versions of the Talbot contour put too much emphasis on the point at infinity. Future investigations will be devoted to this idea of using only a section of the Talbot contour, as well as the use of alternate contours (a few of which will be discussed in §6).

In the next section we solve two parabolic problems to demonstrate that the error estimates of this section are not only applicable to the model problem (2.1), but also to the solution of PDEs.

5 Application to PDEs

The prototype parabolic PDE is of course the heat equation, and no set of tests would be complete without it. We therefore consider

$$u_t = u_{xx}, \quad 0 \leq x \leq \pi, \quad (5.1)$$

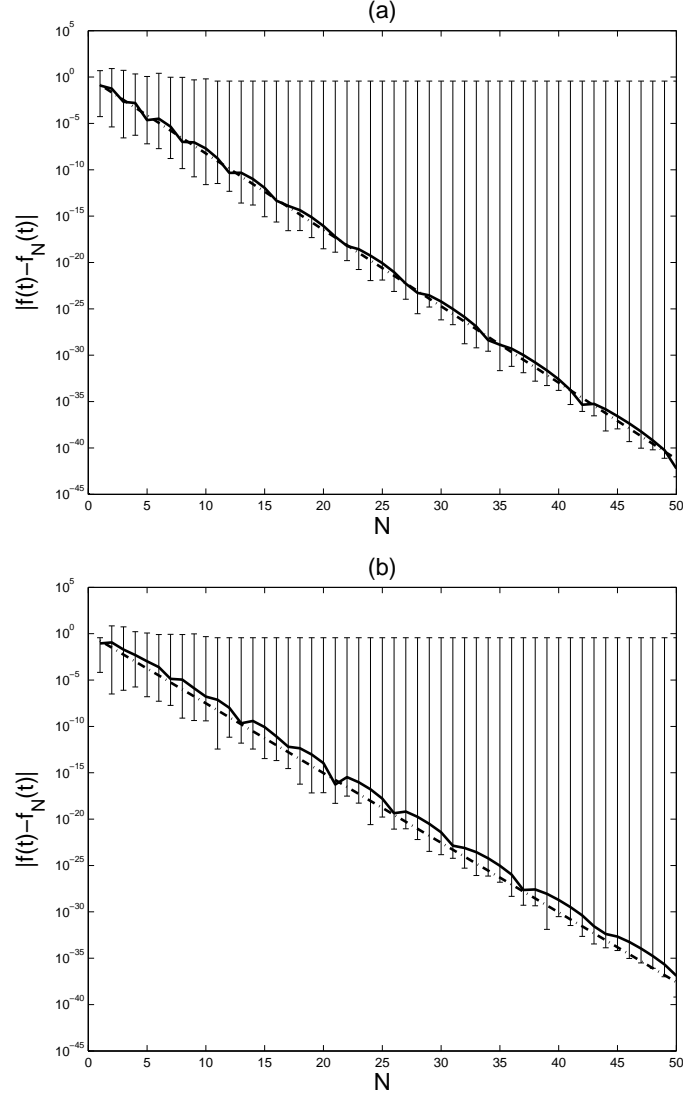


Figure 6: Absolute errors when Talbot's method is applied to the model function (1.10), using (a) the contour (1.3) and (b) the contour (1.5). The thicker curves represent the errors computed with the optimal parameters (4.6) and (4.2). The dash-dot lines (hardly visible in (a)) represent the theoretical error estimates (4.5) and (4.4). The vertical "line segments" represent the range of errors computed in a uniform sampling of parameter space $(s, m, n) \in (0, 1)^3$. Other parameter values were $t = 1, \lambda = -1$.

subject to boundary conditions

$$u(0, t) = 0, \quad u(\pi, t) = 1, \quad t > 0, \quad (5.2)$$

and an initial condition

$$u(x, 0) = 0, \quad 0 \leq x \leq \pi. \quad (5.3)$$

The exact solution can be represented as a Fourier sine series [2, p. 91], which is the standard textbook solution but slowly convergent for small t . An alternative solution, more efficient for small t , is an infinite series involving the complementary error function [2, p. 93].

For numerical implementation, we define $v(x, t) = u(x, t) - x/\pi$ and rewrite the problem as

$$v_t = v_{xx}, \quad (5.4)$$

now with homogeneous boundary conditions

$$v(0, t) = 0, \quad v(\pi, t) = 0, \quad t > 0 \quad (5.5)$$

but inhomogeneous initial condition

$$v(x, 0) = -x/\pi, \quad 0 \leq x \leq \pi.$$

To semi-discretize (5.4), a suitable partition $\{x_j\}_{j=1}^M$ of $[0, \pi]$ is introduced, along with an $M \times M$ matrix D that represents the approximation to d^2/dx^2 and which incorporates the boundary conditions (5.5). The approximation to (5.1) is then given by the linear system of ODEs

$$\mathbf{v}_t = D\mathbf{v}, \quad \mathbf{v}(0) = \mathbf{v}_0. \quad (5.6)$$

Here $\mathbf{v} = \mathbf{v}(t)$ is the $M \times 1$ column vector $[v_1(t), v_2(t), \dots, v_M(t)]^T$, with $v_j(t)$ representing the approximation to $v(x_j, t)$. Likewise \mathbf{v}_0 is the vector that consists of samples of $v(x, 0)$ at the grid-points x_j .

Traditionally, the system (5.6) is solved by the method-of-lines, i.e., a multistep or Runge-Kutta formula would be used to compute approximations to the systems of ODEs. Here we adopt instead the transform approach outlined in §1. That is, we compute the approximation to $\mathbf{v}(t)$ by the midpoint sum

$$\mathbf{v}_N(t) = \frac{1}{N} \operatorname{Im} \left\{ \sum_{k=0}^{N-1} e^{z(\theta_k)t} z'(\theta_k) \mathbf{V}_k \right\}, \quad (5.7)$$

where $z(\theta)$ is given by (1.3) (or (1.5)) and θ_k by (1.7). The vectors \mathbf{V}_k are solved from

$$\left(z(\theta_k) I - D \right) \mathbf{V}_k = \mathbf{v}_0, \quad k = 0, \dots, N-1. \quad (5.8)$$

The details of our particular implementation are as follows. Since we have established that the Talbot contour (1.3) is superior to the contour (1.5), we consider only the former. As for the choice of $\{x_j\}_{j=1}^M$ and D , we use the Chebyshev spectral collocation method,

i.e., the nodes are the Chebyshev points of the second kind, and D is the corresponding spectral second derivative matrix incorporating the boundary conditions (5.2). Note that the canonical interval for the Chebyshev points is $[-1, 1]$, which we transform to $[0, \pi]$ with a linear transformation $x \mapsto (\pi/2)(x+1)$. (Codes for computing $\{x_j\}_{j=1}^M$ and D and further implementation details of the spectral method may be found in [9, 21, 28].)

We shall report errors in the L_2 -norm, which we approximate by the Clenshaw-Curtis rule. This quadrature rule is the natural companion to the Chebyshev collocation method, as the two sets of nodes are identical. That is, we define as error norm

$$E_N(t) = \sqrt{\frac{\pi}{2} \sum_{j=1}^M w_j \left(v(x_j, t) - v_j(t) \right)^2}, \quad (5.9)$$

where the w_j are the weights defined in [21, p. 128], and the factor $\pi/2$ comes from the transformation of the interval $[-1, 1]$ to $[0, \pi]$. The exact solution, $v(x, t)$, was computed by the series expansions alluded to below (5.3).

Our first aim is to demonstrate that the convergence estimate (4.5), derived for the model problem (2.1), is also valid for the solution of a PDE. In the latter case, there is of course a whole spectrum of λ 's present, not only the single λ that was assumed in §2 and §3. For this reason we chose the side conditions (5.2)–(5.3) to represent a discontinuous solution at $t = 0$. Our interest will therefore be in the regime $t \rightarrow 0$, when high frequency modes are relevant.

In Figure 7 we show solutions of (5.1)–(5.3) at various values of t . We also show the error, $E_N(t)$, as a function of N , for the corresponding values of t . We have chosen the $M \times M$ Chebyshev matrices D sufficiently large to fully resolve the solution, i.e., the errors reported in the figure are solely due to the Laplace Transform quadrature error and not due to inadequate spatial resolution. Owing to the smoothing property of the heat equation the order of D can of course be reduced as t increases, and suitably large values of M were determined by trial-and-error.

Assessing these figures, it is clear that the error estimate (4.5) is valid for this problem, even for small t . In addition, one should keep in mind that these results can be achieved by solving effectively only $3N/4$ linear systems; recall the discussion at the end of §4. This allows us to formulate the rule-of-thumb stated in the abstract. Suppose an accuracy of $10^{-\ell}$ is required at a particular value of t . By considering

$$e^{-1.8975N/0.75} = 10^{-\ell} \quad \implies \quad N \approx 0.91\ell$$

one concludes that this should require no more than ℓ solutions of the system (5.8).

A valid point of criticism against the rescaling of parameters that we have used, is the fact that the contour becomes dependent on the value of t ; cf. (3.1). This means, e.g., that to solve the problem at the three different values of t shown in Figure 7, three different contours had to be used, and the linear systems (5.8) had to be solved three times. This is not very efficient, but by looking at Figure 8, one sees that the situation is not as bad as it may seem.

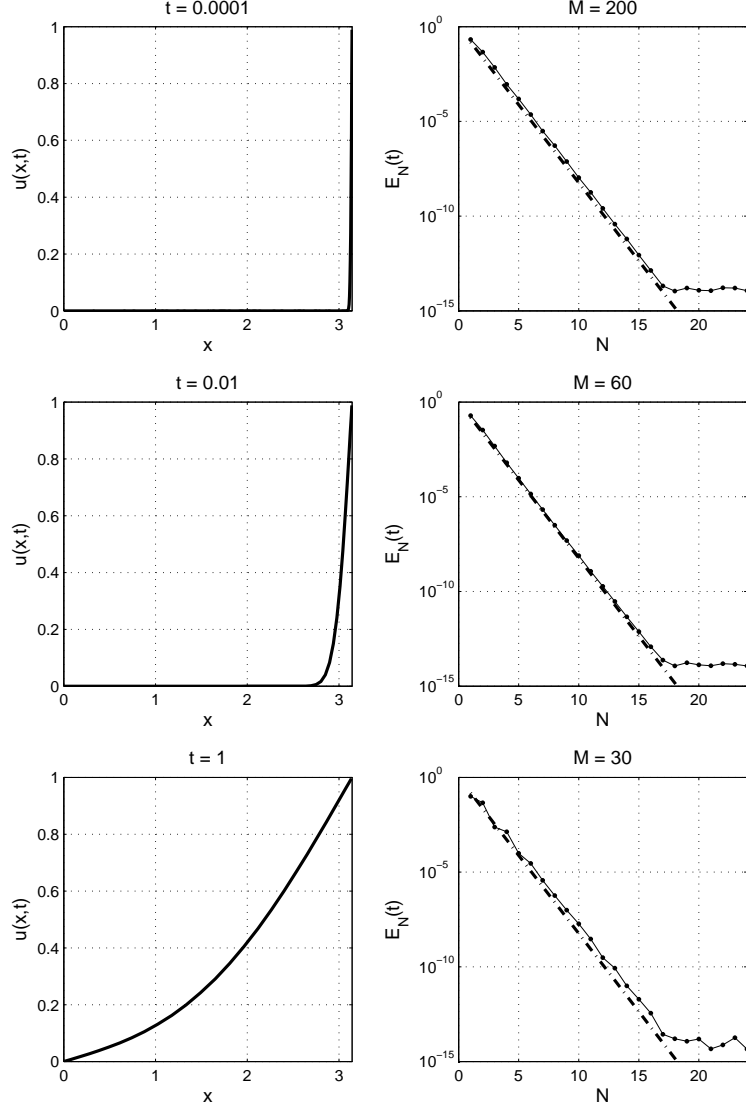


Figure 7: The left column of figures shows the actual solution of (5.1)–(5.3) at various times, computed using the series expansions alluded to below (5.3). The right column shows convergence curves when the solution on the left is approximated with a Chebyshev spectral differentiation matrix of order $M \times M$, and Talbot’s quadrature rule (5.7). The latter was implemented using $2N$ subintervals on the contour (1.3) using optimal parameters (4.6). The thinner, dotted curves show the computed errors, and the thicker, dashdot lines the error model $\exp(-1.8975N)$; cf. (4.5). The error norm, $E_N(t)$, is defined by (5.9).

In that figure, we have computed the errors as functions of t . Instead of varying the parameters with t , as in (3.1), we have fixed the parameters corresponding to a value $t = \tau$. Using the values of (s, m, n) in (4.6), and $N = 20$, this means the values of the contour parameters were approximately

$$\sigma = -481.4, \quad \mu = 644.3 \quad \nu = 0.5653, \quad (5.10)$$

for the error curve on the left ($\tau = 0.02$),

$$\sigma = -96.28, \quad \mu = 128.9, \quad \nu = 0.5653,$$

for the error curve in the middle ($\tau = 0.1$), and

$$\sigma = -17.51, \quad \mu = 23.43, \quad \nu = 0.5653, \quad (5.11)$$

for the error curve on the right ($\tau = 0.55$).

The conclusion to be drawn from this figure is that if the problem (5.1)–(5.3) has to be solved on the interval $[10^{-2}, 1]$ to an accuracy 10^{-10} , say, it can be done using only these three contours. With $M = 50$ and $N = 20$ it means solving the 50×50 linear system (5.8) only $3 \times 20 = 60$ times.

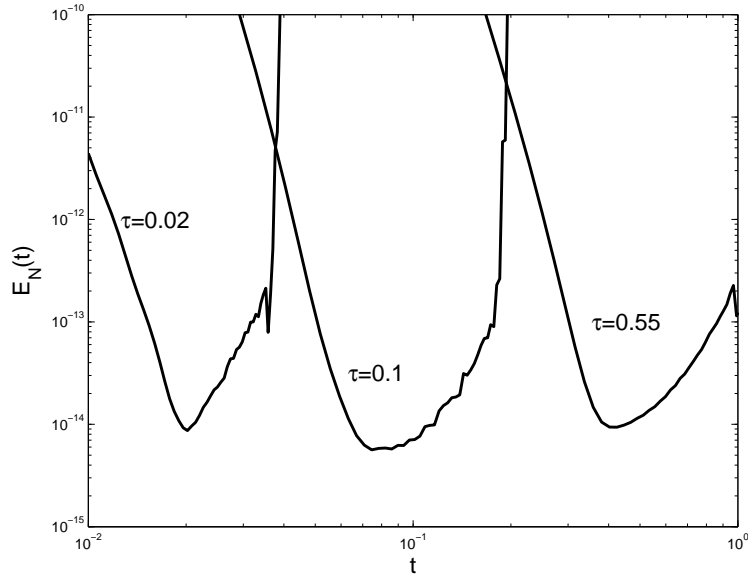


Figure 8: Error curves when solving the problem (5.1)–(5.3) using a Chebyshev differentiation matrix of order 50×50 . The Talbot quadrature (5.7) was used on three contours of the form (1.3), using $N = 20$. The three sets of contour parameters (σ, μ, ν) are defined by (5.10)–(5.11).

As a second example, we consider the fractional heat equation

$$D_t^\alpha u = u_{xx}, \quad (5.12)$$

subject to homogeneous Dirichlet boundary conditions of the form (5.5), and initial condition

$$u(x, 0) = \sin x, \quad 0 \leq x \leq \pi. \quad (5.13)$$

Here D_t^α is the Caputo fractional derivative, defined for continuously differentiable functions $f(t)$ by [17, p. 79]

$$D_t^\alpha f(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{f'(s)}{(t-s)^\alpha} ds, \quad (0 < \alpha < 1).$$

It can be shown [17, p. 79] that if $f(t)$ is twice continuously differentiable, then in the limit $\alpha \rightarrow 1$ this formula reproduces the ordinary derivative, in which case (5.12) reduces to the standard heat equation (5.1).

The analytical solution to (5.12)–(5.13) can be written as

$$u(x, t) = M(t) \sin x,$$

where $M(t)$ can be expressed in terms of the Mittag-Leffler function. In the case $\alpha \rightarrow 1$, it reduces of course to $M(t) = e^{-t}$. In the case $\alpha = 1/2$, the function can be expressed in terms of the complementary error function, namely

$$M(t) = e^t \operatorname{erfc}(\sqrt{t}).$$

The qualitative properties of this $\alpha = 1/2$ solution are similar to those of the ordinary heat equation, but steady-state is approached on a longer time scale (a property referred to in the literature as sub-diffusion).

For the numerical solution of (5.12)–(5.13), one takes a Laplace transform of (5.12), and finds that the definition of $F(z)$ in the Bromwich integral (1.2) needs to be modified to

$$F(z) = (zI - z^{1/2}D)^{-1}.$$

Here D is again any discrete approximation to the second derivative operator with homogeneous Dirichlet boundary conditions on $[0, \pi]$; we shall continue to use a Chebyshev differentiation matrix of order $M \times M$. The modification to the Talbot method (5.7)–(5.8) is obvious: all it requires is that the scalar $z(\theta_k)^{1/2}$ be inserted to multiply D in (5.8).

Finding optimal parameters for Talbot's method for the problem (5.12)–(5.13) would mean analyzing $F(z) = 1/(z - z^{1/2}\lambda)$ as test function as we analyzed $F(z) = 1/(z - \lambda)$ in §3. Note that the singularities are no longer isolated, but a branch cut that covers the negative real axis. Instead of performing such an analysis, we shall merely demonstrate numerically that Talbot's method, with the rescaling of parameters suggested in §3, is very accurate for this problem as well. The error curves shown in Figure 9 confirm that the convergence rate is, to a good approximation, once again given by $E_N(t) = O(e^{-1.8975N})$.

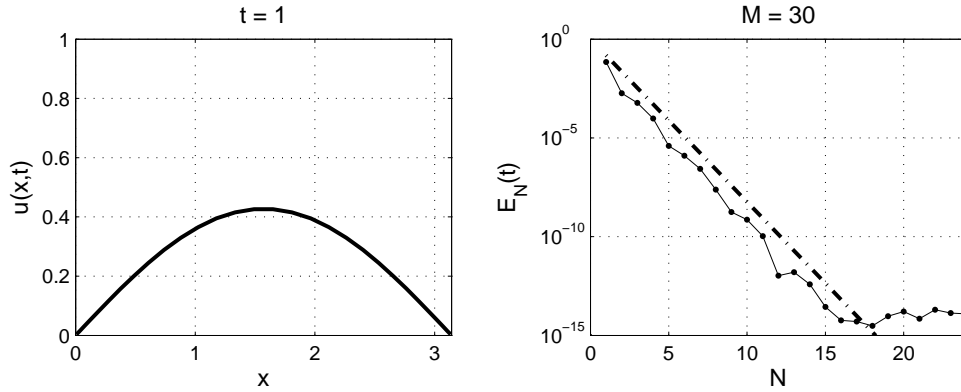


Figure 9: Same as Fig. 7, but the problem is the fractional heat equation, (5.12)–(5.13), with $\alpha = 1/2$.

6 Comparisons

In this section we draw a few comparisons with related work in the literature. First, we contrast Talbot’s contour with a few other contours. Second, we compare our new parameter choices with similar suggestions made by Talbot in [20].

The contours used in [5, 10, 14, 15, 18, 19] are all of the form

$$z = z(w), \quad -\infty < w < \infty,$$

with the property that $\operatorname{Re} z \rightarrow -\infty$ as $w \rightarrow \pm\infty$. With such a change of variables, the integral (1.2) can be expressed as

$$f(t) = \frac{1}{2\pi i} \int_{-\infty}^{\infty} e^{z(w)t} \mathbf{F}(z(w)) z'(w) dw.$$

The trapezoidal rule is applied to this integral, on a grid

$$w_k = kh, \quad k = -N, \dots, N.$$

The step size, h , is chosen according to some grid refinement strategy $h = h(N)$, where $h \rightarrow 0$ as $N \rightarrow \infty$.

Table 2 is a summary of various contours found in the literature, with grid refinement strategy $h = h(N)$, and expected convergence rate. The parameters α , μ , ν and σ are all positive, and c is a generic positive constant. The first and third contours are hyperbolas, and the second a parabola; see Figure 10.

When applied to the model problem (1.10), with $\lambda = -1$ and $t = 1$, each of the grids shown in Figure 10 produces an absolute error less than 10^{-5} . (It might be noted that the method based on contour Γ_1 was further improved in [15, 19], by making another variable transformation in the integral that stretches the outlying nodes further outward along the contour.)

Table 2: Summary of various contours used in literature.

$z = z(w)$	$h = h(N)$	Conv. Rate	Ref.
$\Gamma_1: z = \sigma - \sqrt{w^2 + \alpha^2} + i\nu w$	$h = c/N^{1/2}$	$\exp(-cN^{1/2})$	[15]
$\Gamma_2: z = \sigma - \mu w^2 + i\nu w$	$h = c/N^{2/3}$	$\exp(-cN^{2/3})$	[10]
$\Gamma_3: z = \mu(1 - \sin(\alpha - iw))$	$h = c(\log N)/N$	$\exp(-cN/\log N)$	[14]

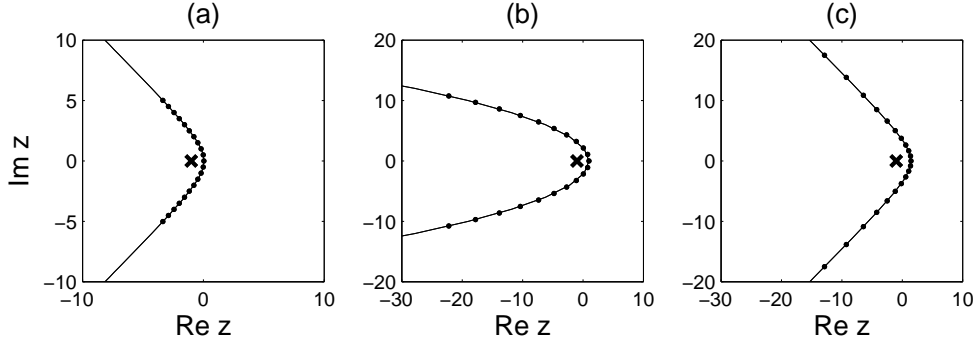


Figure 10: The three contours of Table 2, with (a) Γ_1 , (b) Γ_2 , and (c) Γ_3 . The dots represent the nodes (with $N = 10$), and the crosses the pole at $z = -1$, which corresponds to $\lambda = -1$ in the model equation (1.10).

For a comparison of convergence rates, we recall from §2 and [20] that Talbot's method, with fixed parameters, converges as $O(e^{-cN^{1/2}})$, $N \rightarrow \infty$. This makes it in theory as accurate as the first method in Table 2, and less accurate than the other two. With the new sets of parameters suggested here, however, we have improved the convergence of Talbot's method to $O(e^{-cN})$, with the value of c fairly large; cf. (4.4) and (4.5). Indeed, the numerical results shown in Figures 6, 7 and 9 indicate accuracy at least as good as the best results presented in [5, 10, 14, 15, 18, 19], and in many cases better, particularly for larger values of N .

We are not excluding the possibility that some of the methods summarized in the above table can be improved by making the parameters dependent on t . If successful, it is entirely possible that one or more of these contours can improve on the convergence rates obtained in this paper. We trust the framework presented in this paper can serve as basis for such investigations.

In this paper we have focused on pure parabolic problems, with eigenvalues on the negative real axis. When the spectrum is no longer on this axis, some of the contours in Table 2 offer distinct advantages over Talbot's. In particular, when the operator is sectorial, i.e., when the spectrum is contained in a wedge shaped region with apex at the origin and with imaginary parts extending to $\pm\infty$, the hyperbolic contours can be

made to enclose the spectrum whereas the Talbot contour cannot. See, for example, the problems solved in [14].

Turning to a comparison with earlier work on Talbot's method, we remark that in [20] some suggestions for parameter selection were made. We quote from the abstract of that paper "The required number of points depends on t , and for moderate t is typically 11 for orders of 10^{-6} , 18 for order 10^{-10} , 35 for order 10^{-20} ." With the convergence rate $O(e^{-1.8975N})$ derived here, these figures can be improved to about 8 points for orders of 10^{-6} , 13 for order 10^{-10} , and 25 for order 10^{-20} . In addition, with the omission of outlying nodes as discussed in the final paragraphs of §4, these numbers can be reduced further to 6, 10, and 19 points for the three accuracies quoted. Note that Talbot did not explicitly report a convergence rate in the form of $O(e^{-cN})$, as the focus of [20] was more on finding parameter sets that will meet given error criteria, namely single and double precision accuracies.

For a quantitative comparison, the parameters suggested in [20] are

$$\sigma = 0, \quad \mu = \frac{\omega}{t}, \quad \nu = 1, \quad (6.1)$$

for the case when all singularities are real and negative. The suggested value of ω is 6 and 11, respectively, for single and double precision tolerances. These values of ω are to be compared with our $mN \approx 0.6443N$, and we find that $\omega = 6$ and 11 correspond to $N = 9$ and 17. This is consistent with our results, as $\exp(-1.8975 \times 9) \approx 3.9 \cdot 10^{-8}$ (approx. single precision) and $\exp(-1.8975 \times 17) \approx 9.8 \cdot 10^{-15}$ (approx. double precision). The values $\sigma = 0$ and $\nu = 1$ cited in (6.1), however, are without a doubt suboptimal. We remark that while some asymptotics were used in [20], Talbot relied largely on empirical and heuristical reasoning to derive the estimate (6.1).

To be fair to Talbot, the aims of the paper [20] were much more ambitious than those of the present paper. To begin with, all singularity distributions were considered, not just poles on the negative imaginary axis. In addition, Talbot tried to find optimal parameters for finite precision tolerances, and therefore had to deal with the locations of the singularities. By contrast, we let $N \rightarrow \infty$, thereby making the errors independent of the singularities, and trusted in the power of asymptotics to make the parameters thus found relevant for finite (indeed, relatively small) values of N as well.

Acknowledgments. This work was done at the Oxford University Computing Laboratory, while the author was a Visiting Fellow at Exeter College. The National Research Foundation in South Africa provided the main funding. The University of Stellenbosch, South Africa, is acknowledged for granting sabbatical leave, as well as providing additional funding. At the Computing Laboratory, Nick Trefethen's hospitality extended from arranging accommodation to making substantial improvements to the content of this paper, truly wonderful support without which this research would not have been completed. Kevin Burrage introduced the author to the fractional heat equation (5.12), and Nick Gould offered many useful suggestions regarding the numerical optimization described in §4. While at Oregon State University, the author and Fusen Lin had many discussions on Talbot's method, and one chapter of Lin's thesis is devoted to this method [13].

References

- [1] Mark J. Ablowitz and Athanassios S. Fokas. *Complex Variables: Introduction and Applications*. Cambridge University Press, Cambridge, second edition, 2003.
- [2] M. Ya. Antimirov, A. A. Kolyshkin, and Rémi Vaillancourt. *Applied Integral Transforms*. American Mathematical Society, Providence, RI, 1993.
- [3] Carl M. Bender and Steven A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. McGraw-Hill Book Co., New York, 1978.
- [4] John P. Boyd. Asymptotic coefficients of Hermite function series. *J. Comput. Phys.*, 54(3):382–410, 1984.
- [5] E. Cuesta and C. Palencia. A numerical method for an integro-differential equation with memory in Banach spaces: qualitative properties. *SIAM J. Numer. Anal.*, 41:1232–1241, 2003.
- [6] Philip J. Davis. On the numerical integration of periodic analytic functions. In *On numerical approximation. Proceedings of a Symposium, Madison, April 21-23, 1958*, Edited by R. E. Langer. The University of Wisconsin Press, Madison, 1959.
- [7] Dean G. Duffy. On the numerical inversion of laplace transforms: comparison of three new methods on characteristic problems from applications. *ACM Trans. Math. Software*, 19:333–359, 1993.
- [8] David Elliott and George Szekeres. Some estimates of the coefficients in the Chebyshev series expansion of a function. *Math. Comp.*, 19:25–32, 1965.
- [9] Bengt Fornberg. *A Practical Guide to Pseudospectral Methods*. Cambridge University Press, Cambridge, 1996.
- [10] Ivan P. Gavrilyuk and Vladimir L. Makarov. Exponentially convergent parallel discretization methods for the first order evolution equations. *Comput. Methods Appl. Math.*, 1:333–355, 2001.
- [11] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, MD, third edition, 1996.

- [12] E. T. Goodwin. The evaluation of integrals of the form $\int_{-\infty}^{\infty} f(x)e^{-x^2}dx$. *Proc. Cambridge Philos. Soc.*, 45:241–245, 1949.
- [13] F. Lin. *Numerical Inversion of Laplace Transforms by Trapezoidal-Type Methods*. Ph. D. Dissertation. Oregon State University, Corvallis, OR, 2003.
- [14] M. López-Fernández and C. Palencia. On the numerical inversion of the Laplace transform of certain holomorphic mappings. *Appl. Numer. Math.*, 51:289–303, 2004.
- [15] William McLean and Vidar Thomée. Time discretization of an evolution equation via Laplace transforms. *IMA J. Numer. Anal.*, 24:439–463, 2004.
- [16] Cleve Moler and Charles Van Loan. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM Rev.*, 45:3–49, 2003.
- [17] Igor Podlubny. *Fractional Differential Equations*. Academic Press Inc., San Diego, CA, 1999.
- [18] Dongwoo Sheen, Ian H. Sloan, and Vidar Thomée. A parallel method for time-discretization of parabolic problems based on contour integral representation and quadrature. *Math. Comp.*, 69:177–195, 2000.
- [19] Dongwoo Sheen, Ian H. Sloan, and Vidar Thomée. A parallel method for time discretization of parabolic equations based on Laplace transformation and quadrature. *IMA J. Numer. Anal.*, 23:269–299, 2003.
- [20] A. Talbot. The accurate numerical inversion of Laplace transforms. *J. Inst. Math. Appl.*, 23:97–120, 1979.
- [21] Lloyd N. Trefethen. *Spectral Methods in MATLAB*. SIAM, Philadelphia, PA, 2000.
- [22] Lloyd N. Trefethen. Private communication. 2005.
- [23] A. M. Turing. A method for the calculation of the zeta-function. *Proc. London Math. Soc. (2)*, 48:180–197, 1943.
- [24] J. M. Varah. On the numerical solution of ill-conditioned linear systems with applications to ill-posed problems. *SIAM J. Numer. Anal.*, 10:257–267, 1973.
- [25] J. A. C. Weideman. Computation of the complex error function. *SIAM J. Numer. Anal.*, 31(5):1497–1518, 1994.
- [26] J. A. C. Weideman. Erratum: “Computation of the complex error function”. *SIAM J. Numer. Anal.*, 32(1):330–331, 1995.
- [27] J. A. C. Weideman. Numerical integration of periodic functions: a few examples. *Amer. Math. Monthly*, 109:21–36, 2002.
- [28] J. A. C. Weideman and S. C. Reddy. A MATLAB differentiation matrix suite. *ACM Trans. Math. Software*, 26:465–519, 2000.

Appendix: Saddle points and critical points corresponding to cases (a)–(d) in Fig. 4.

Case	Parameters	Roots of (3.6)(+)	Roots of (3.6)(−)	Roots of (3.7)	$\theta_+, \theta_-, \theta_*$	d_+, d_-, d_*	Conv. Rate
(a)	$s = 0.8$ $m = 1.05$ $n = 0.4$	$\pm 3.9751 - 0.8123i$ $\pm 2.2930 + 0.8123i$	$\pm 4.1761 + 0.9975i$ $\pm 2.0710 - 0.9975i$	$2.8686i$ $\pm 0.4351 + 1.3633i$	$-2.2930 + 0.8123i$ $-2.0710 - 0.9975i$ $-0.4351 + 1.3633i$	-2.2380 -1.4859 -2.7267	$\exp(-1.4859N)$
(b)	$s = 0.6$ $m = 0.8$ $n = 0.5$	$\pm 3.8710 - 0.7145i$ $\pm 2.4033 + 0.7145i$	$\pm 4.0372 + 0.8700i$ $\pm 2.2258 - 0.8700i$	$0.6780i$ $\pm 1.9043 + 1.9110i$	$-2.4033 + 0.7145i$ $-2.2258 - 0.8700i$ $0.6780i$	-2.1185 -1.4908 -1.3560	$\exp(-1.3560N)$
(c)	$s = 0.7556$ $m = 0.8597$ $n = 0.3029$	$\pm 3.9213 - 0.7620i$ $\pm 2.3503 + 0.7620i$	$\pm 4.0319 + 0.8652i$ $\pm 2.2315 - 0.8652i$	$5.2713i$ $0.8652i$ $0.8652i$	$-2.3503 + 0.7620i$ $-2.2315 - 0.8652i$ $0.8652i$	-2.1524 -1.7303 -1.7303	$\exp(-1.7303N)$
(d)	$s = 0.8$ $m = 1.05$ $n = 2$	$\pm 3.7800 - 0.6281i$ $\pm 2.4979 + 0.6281i$	$\pm 7.2117 - 3.3570i$ $6.4764i$ $0.2377i$	$0.1205i$ $\pm 3.0822 + 0.4993i$	N/A	N/A	N/A