THE QUADRATIC HERMITE-PADÉ APPROXIMATION

RICHARD G. BROOKES

A thesis presented for the degree of
Doctor of Philosophy in Mathematics
in the University of Canterbury

University of Canterbury
1989

PHYSICAL
SCIENCES
LIBRARY
THESIS
copy 2

Contents

Abstract

This thesis is concerned with the existence, behaviour and performance of the quadratic Hermite-Padé approximation.

It starts with the definition of the general Hermite-Padé approximation. Some of the problems which arise, particularly those of finding Hermite-Padé forms and the existence of approximations are discussed. Chapter 3 solves the existence problem in the quadratic case whilst Chapter 2 presents a recurrence algorithm for finding quadratic forms which can easily be extended to general Hermite-Padé forms.

Chapters 4 and 5 compare the performance of the quadratic, Padé and Taylor approximations using particular examples over a variety of regions. Many graphs and contour maps of the various approximations and error functions are given. The quadratic approximation is shown to be superior in these cases.

Finally, in Chapter 6, a theorem concerning sequences of quadratic approximations is presented and the structure of the quadratic table is explored.

Acknowledgements

I would like to thank my supervisor Allan W. McInnes for all his help and guidance and Ann Tindall for a great deal of difficult typing.

CHAPTER 1

HERMITE-PADÉ APPROXIMATIONS

1. Introduction

This thesis will concentrate largely on the characteristics of the quadratic Hermite-Padé approximation. Firstly, however, it is necessary to understand the formulation of the more general Hermite-Padé approximation, some of its elementary properties and some of its problems.

2. The general Hermite-Padé approximation

(i) Let $f(x)$ be a function, analytic in a neighbourhood of the origin whose power series expansion about the origin is known.

(ii) Let $g_0, g_1, \ldots, g_n$ be functions such that $\forall i \in \{0, \ldots, n\}$ $g_i(f(x))$ is analytic in a neighbourhood of the origin and its power series expansion about the origin is known.

(iii) Let $A_0, A_1, \ldots, A_n \in \mathbf{Z}^+ \cup \{-1\}$ and $N = \sum_{i=0}^{n} A_i$.

(iv) Let $a_0(x), a_1(x), \ldots, a_n(x)$ be polynomials in x with $\deg(a_i(x)) \leq A_i$ $\forall i \in \{0, \ldots, n\}$, such that

$$\sum_{i=0}^{n} a_i(x)\, g_i(f(x)) = O\left(x^{N+n}\right) \tag{1}$$

(The only polynomial of degree -1 is, by definition, the zero polynomial.)

Note that such $a_i(x)$, not all zero, must exist since (1) represents a homogeneous system of $N+n$ linear equations in the $N+n+1$ unknown coefficients of the $a_i(x)$.

A set of $a_i(x)$ derived in this way is known as a $(A_n, A_{n-1}, \ldots, A_0)$ Hermite-Padé form for the system $g_i(f(x))$. Such a form will be represented as $\{a_n(x), ..., a_0(x)\}$ or as $\sum_{i=0}^{n} a_i(x) g_i(y)$.

Set

$$\sum_{i=0}^{n} a_i(x)\, g_i(y(x)) = 0. \tag{2}$$

It is the solution of (2) for $y(x)$ (not necessarily unique) that gives a Hermite-Padé approximation for $f(x)$.

This is the standard way of defining Hermite-Padé approximations. See, for example, Della Dora and Di Crescenzo [5], Paszkowski [8] and Baker and Lubinsky [2] in which there appears an extensive bibliography.

Examples

(i) Setting $g_i(f(x)) = \dfrac{d^i f(x)}{dx^i}$ gives the differential Hermite-Padé approximants (see [2]). The case $n = 1$ is that of the Baker $D-\log$ approximation (see [1]).

(ii) Setting $g_i(f(x)) = f(x)^i$ gives the algebraic Hermite-Padé approximants of which there are several important special cases:

1. $n = 1, A_1 = 0$ gives the Taylor approximation.
2. $n = 1$ gives the Padé approximation.
3. $n = 2$ gives the quadratic approximation. (again, see [2] and the references therein)

3. Finding a Hermite-Padé form

Let $g_i(f(x)) = \sum_{j=0}^{\infty} g_i^j x^j, \quad a_i(x) = \sum_{j=0}^{A_i} a_i^j x^j.$

Then the system of linear equations given by (1) can be represented as the following matrix equation:

$$\begin{bmatrix} g_0^0 & 0 & 0 & & & g_n^0 & 0 & 0 & \\ g_0^1 & g_0^0 & 0 & & & g_n^1 & & 0 & \\ g_0^2 & g_0^1 & & & & & & & \\ & & & g_0^0 & & & & g_n^0 & \\ \vdots & \vdots & \vdots & \vdots & \cdots\cdots & \vdots & \vdots & \vdots & \vdots \\ & & & & & & & & \\ & & & & & & & & \\ g_0^{N+n-1} & g_0^{N+n-2} & \cdots & g^{N+n-A_0-1} & & g_n^{N+n-1} & & g^{N+n-A_n-1} & \end{bmatrix} \begin{bmatrix} a_0^0 \\ a_0^1 \\ \vdots \\ a_0^{A_0} \\ \vdots \\ \\ \\ a_k^{A_k} \end{bmatrix} = 0 \qquad (3)$$

So, to obtain the possible $(A_n, \ldots, A_0)$ Hermite-Padé forms it is sufficient to row-reduce the matrix in (3) to find the linear subspace of solutions. Of course there are an infinite number of elements in this subspace, which raises the questions of normalisation and precisely which solution it is best to choose. Certainly forms which are linear multiples of each other yield the same solutions for $y(x)$ so normalisation is largely irrelevant. However, the question of which element to choose from a multi-dimensional solution space is more difficult to answer and will be investigated in Chapter 6.

3.1 Numerical Difficulties

The problem with this simple approach is that attempts to implement it as a practical algorithm are subject to extreme numerical difficulties. Standard linear equation solvers applied to (3) go hopelessly astray. These problems are best illustrated by the following table of approximate singular values for several example matrices. The solution space for Padé-Hermite

forms is in each case one dimensional so that each of these matrices has one singular value of zero, but unfortunately has several other singular values which are very small. This leads one to expect that numerical attempts to diagonalise these matrices will yield inaccurate results. This has been verified in practice.

Approximate singular values of the matrix for the $(8,8,8)$ quadratic form.

$f(x) = e^x$	$f(x) = \log(1+x)$	$f(x) = \sin x$
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^0	10^0	10^0
10^{-1}	10^0	10^0
10^{-2}	10^{-1}	10^{-1}
10^{-3}	10^{-2}	10^{-2}
10^{-4}	10^{-3}	10^{-3}
10^{-6}	10^{-4}	10^{-4}
10^{-8}	10^{-5}	10^{-5}
10^{-9}	10^{-6}	10^{-7}
10^{-11}	10^{-7}	10^{-8}
10^{-13}	10^{-9}	10^{-9}
10^{-15}	10^{-10}	10^{-11}
10^{-17}	10^{-11}	10^{-13}
10^{-19}	10^{-13}	10^{-15}
10^{-21}	10^{-15}	10^{-15}
10^{-22}	10^{-16}	10^{-17}
10^{-23}	10^{-17}	10^{-20}
10^{-25}	10^{-18}	10^{-23}
0	0	0

Other approaches to the problem such as recursive algorithms (for example that developed by Loi and McInnes [7] or that outlined in Chapter 2) appear to suffer from similar problems.

3.2 Symbolic Reduction

A partial solution to these difficulties is to use one of the symbolic manipulation packages such as MACSYMA. If $g_i^j \in \mathbf{Q} \quad \forall i,j$ then it is a relatively simple task to write a program which will return an exact basis for the solution space of (3). This makes it easy to generate tables of Hermite-Padé forms for many functions such as $e^x, \log(1+x), \cos(x), \sqrt[3]{1+x}$ etc. All the exact Padé-Hermite forms appearing in this thesis have been found in this way.

The capability to produce these exact forms is extremely important. All the subsequent work has been motivated and guided by these examples, many of which do not appear to have

been published (although Borwein [4] has some specific results concerning quadratic forms for $e^x, \log(x), x^{\frac{1}{n}}$). Following are tables for the quadratic forms $\{(i,j,k) : i,j,k \in \{0,1,2\}\}$ for the functions $\log(1+x), \cos(x), \sqrt[3]{1+x}$. The spaces of quadratic forms for $\log(1+x)$ and $\sqrt[3]{1+x}$ are one-dimensional, however those for $\cos(x)$ are often two-dimensional.

Table 1: Hermite-Padé forms for $\log(1+x)$

$A_2 = 0$

A_0 \ A_1	0	1	2
0	$a_0(x) = 0$ $a_1(x) = 0$ $a_2(x) = 1$	$a_0(x) = 0$ $a_1(x) = -x$ $a_2(x) = 1$	$a_0(x) = 0$ $a_1(x) = x^2 - 2x$ $a_2(x) = 2$
1	$a_0(x) = -2x$ $a_1(x) = 2$ $a_2(x) = 1$	$a_0(x) = -6x$ $a_1(x) = 2x + 6$ $a_2(x) = 1$	$a_0(x) = -48x$ $a_1(x) = -x^2 + 18x + 48$ $a_2(x) = 6$
2	$a_0(x) = x^2 - 6x$ $a_1(x) = 6$ $a_2(x) = 2$	$a_0(x) = -x^2 - 18x$ $a_1(x) = 8x + 18$ $a_2(x) = 2$	$a_0(x) = -7x^2 - 30x$ $a_1(x) = 2x^2 + 20x + 30$ $a_2(x) = 2$

$A_2 = 1$

A_0 \ A_1	0	1	2
0	$a_0(x) = 0$ $a_1(x) = 0$ $a_2(x) = x$	$a_0(x) = 0$ $a_1(x) = -2x$ $a_2(x) = x + 2$	$a_0(x) = 0$ $a_1(x) = -x^2 - 6x$ $a_2(x) = 4x + 6$
1	$a_0(x) = -6x$ $a_1(x) = 6$ $a_2(x) = x + 3$	$a_0(x) = 12x$ $a_1(x) = -6x - 12$ $a_2(x) = x$	$a_0(x) = 120x$ $a_1(x) = -x^2 - 66x - 120$ $a_2(x) = 14x + 6$
2	$a_0(x) = -3x^2 - 6x$ $a_1(x) = 6$ $a_2(x) = 4x + 6$	$a_0(x) = -x^2 + 30x$ $a_1(x) = -16x - 30$ $a_2(x) = 4x + 2$	$a_0(x) = -9x^2 + 30x$ $a_1(x) = 2x^2 - 12x - 30$ $a_2(x) = 8x + 6$

$A_2 = 2$

A_0 \ A_1	0	1	2
0	$a_0(x) = 0$ $a_1(x) = 0$ $a_2(x) = x^2$	$a_0(x) = 0$ $a_1(x) = 12x$ $a_2(x) = x^2 - 6x - 12$	$a_0(x) = 0$ $a_1(x) = -3x^2 - 6x$ $a_2(x) = x^2 + 6x + 6$
1	$a_0(x) = 24x$ $a_1(x) = -24$ $a_2(x) = x^2 - 4x - 12$	$a_0(x) = -360x$ $a_1(x) = 192x + 360$ $a_2(x) = x^2 - 36x - 12$	$a_0(x) = 240x$ $a_1(x) = -11x^2 - 150x - 240$ $a_2(x) = 3x^2 + 46x + 30$
2	$a_0(x) = -12x^2$ $a_1(x) = 0$ $a_2(x) = x^2 + 12x + 12$	$a_0(x) = -33x^2 + 270x$ $a_1(x) = -144x - 270$ $a_2(x) = 2x^2 + 60x + 42$	$a_0(x) = 24x^2$ $a_1(x) = -9x^2 - 18x$ $a_2(x) = x^2 - 6x - 6$

Table 2: Hermite-Padé forms for $\cos x$

$A_2 = 0$

A_0 \ A_1	0	1	2
0	$a_0(x) = -1$ $a_1(x) = 1$ $a_2(x) = 0$	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = -2$ $a_1(x) = x^2 + 2$ $a_2(x) = 0$
	$a_0(x) = -1$ $a_1(x) = 0$ $a_2(x) = 1$		$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$
1	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = -1$ $a_1(x) = 3x^2 - 4$ $a_2(x) = 5$
2	$a_0(x) = x^2 - 2$ $a_1(x) = 2$ $a_2(x) = 0$	$a_0(x) = -3x^2 + 7$ $a_1(x) = -8$ $a_2(x) = 1$	$a_0(x) = 5x^2 - 12$ $a_1(x) = x^2 + 12$ $a_2(x) = 0$
	$a_0(x) = x^2 - 1$ $a_1(x) = 0$ $a_2(x) = 1$		$a_0(x) = -3x^2 + 7$ $a_1(x) = -8$ $a_2(x) = 1$

$A_2 = 1$

A_0 \ A_1	0	1	2
0	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = -1$ $a_1(x) = 3x^2 - 4$ $a_2(x) = 5$
1	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = x$ $a_1(x) = -2x$ $a_2(x) = x$	$a_0(x) = -1$ $a_1(x) = 3x^2 - 4$ $a_2(x) = 5$
2	$a_0(x) = -3x^2 + 7$ $a_1(x) = -8$ $a_2(x) = 1$	$a_0(x) = -3x^2 + 7$ $a_1(x) = -8$ $a_2(x) = 1$	$a_0(x) = 11x^2 - 27$ $a_1(x) = 4x^2 + 24$ $a_2(x) = 3$

$A_2 = 2$

A_0 \ A_1	0	1	2
0	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = 5$ $a_1(x) = -16$ $a_2(x) = 3x^2 + 11$	$a_0(x) = -1$ $a_1(x) = 3x^2 - 4$ $a_2(x) = 5$
	$a_0(x) = -2$ $a_1(x) = 2$ $a_2(x) = x^2$		$a_0(x) = 12$ $a_1(x) = -11x^2 - 12$ $a_2(x) = 5x^2$
1	$a_0(x) = 5$ $a_1(x) = -16$ $a_2(x) = 3x^2 + 11$	$a_0(x) = 5$ $a_1(x) = -16$ $a_2(x) = 3x^2 + 11$	$a_0(x) = -3$ $a_1(x) = 64x^2 - 144$ $a_2(x) = 11x^2 + 147$
2	$a_0(x) = -3x^2 + 7$ $a_1(x) = -8$ $a_2(x) = 1$	$a_0(x) = -16x^2 + 39$ $a_1(x) = -48$ $a_2(x) = x^2 + 9$	$a_0(x) = 11x^2 - 27$ $a_1(x) = 4x^2 + 24$ $a_2(x) = 3$
	$a_0(x) = 11x^2 - 24$ $a_1(x) = 24$ $a_2(x) = x^2$		$a_0(x) = -49x^2 + 120$ $a_1(x) = -12x^2 - 120$ $a_2(x) = x^2$

Table 3: Hermite-Padé forms for $\sqrt[3]{1+x}$

$A_2 = 0$

A_0 \ A_1	0	1	2
0	$a_0(x) = 1$ $a_1(x) = -2$ $a_2(x) = 1$	$a_0(x) = 3$ $a_1(x) = -x - 9$ $a_2(x) = 6$	$a_0(x) = 45$ $a_1(x) = 4x^2 - 27x - 171$ $a_2(x) = 126$
1	$a_0(x) = -x$ $a_1(x) = -3$ $a_2(x) = 3$	$a_0(x) = -4x - 3$ $a_1(x) = x - 3$ $a_2(x) = 6$	$a_0(x) = -105x - 90$ $a_1(x) = -x^2 + 33x - 36$ $a_2(x) = 126$
2	$a_0(x) = x^2 - 18x - 9$ $a_1(x) = -36$ $a_2(x) = 45$	$a_0(x) = -x^2 - 42x - 36$ $a_1(x) = 15x - 9$ $a_2(x) = 45$	$a_0(x) = -7x^2 - 84x - 72$ $a_1(x) = 2x^2 + 39x + 9$ $a_2(x) = 63$

$A_2 = 1$

A_0 \ A_1	0	1	2
0	$a_0(x) = -6$ $a_1(x) = 15$ $a_2(x) = x - 9$	$a_0(x) = 3$ $a_1(x) = -5x - 15$ $a_2(x) = 2x + 12$	$a_0(x) = 9$ $a_1(x) = -2x^2 - 39x - 72$ $a_2(x) = 21x + 63$
1	$a_0(x) = -10x - 6$ $a_1(x) = -15$ $a_2 = x + 21$	$a_0(x) = 10x + 9$ $a_1(x) = 5x$ $a_2(x) = x - 9$	$a_0(x) = 105x + 99$ $a_1(x) = -x^2 - 72x - 36$ $a_2(x) = 21x - 63$
2	$a_0(x) = -x^2 - 12x - 9$ $a_1(x) = -9$ $a_2(x) = 3x + 18$	$a_0(x) = -x^2 + 18x + 18$ $a_1(x) = -15x - 9$ $a_2(x) = 6x - 9$	$a_0(x) = -7x^2 + 21x + 27$ $a_1(x) = x^2 - 33x - 27$ $a_2(x) = 21x$

$A_2 = 2$

A_0 \ A_1	0	1	2
0	$a_0 = 63$ $a_1(x) = -180$ $a_2(x) = 5x^2 - 18x + 117$	$a_0(x) = -63$ $a_1(x) = 210x + 450$ $a_2(x) = 5x^2 - 102x - 387$	$a_0(x) = 9$ $a_1(x) = -16x^2 - 102x - 126$ $a_2(x) = 5x^2 + 66x + 117$
1	$a_0(x) = 84x + 63$ $a_1(x) = 90$ $a_2(x) = x^2 - 12x - 153$	$a_0(x) = -336x - 315$ $a_1(x) = 210x + 90$ $a_2(x) = x^2 - 54x + 225$	$a_0(x) = 84x + 81$ $a_1(x) = -4x^2 - 78x - 54$ $a_2(x) = x^2 + 30x - 27$
2	$a_0(x) = -14x^2 - 84x - 63$ $a_1(x) = -36$ $a_2(x) = x^2 + 30x + 99$	$a_0(x) = -28x^2 + 168x + 189$ $a_1(x) = -210x - 162$ $a_2(x) = x^2 + 114x - 27$	$a_0(x) = 28x^2 - 27$ $a_1(x) = -8x^2 + 54x + 54$ $a_2(x) = x^2 - 54x - 27$

Unfortunately, this method is heavily dependent on the knowledge of exact g_i^j. It has been verified with various examples that a small perturbation in the values of the f^j gives a large variation of the (quadratic) Hermite-Padé form.

4. Basic properties of the algebraic Hermite-Padé approximants

Attention is now restricted to algebraic approximants, that is, those corresponding to $g_i(f(x)) = f(x)^i$. Some of their basic properties and difficulties will be explored. Many of these difficulties can be solved in the quadratic case.

In the well known case of Padé approximation (see [1]) it is clear that one valid interpretation of the procedure is that one is finding a rational function of the form $\frac{a_0(x)}{a_1(x)}$ (where the $a_i(x)$ are polynomials of degree A_i) with a power series expansion whose first $A_0 + A_1 + 1$ coefficients match those of $f(x)$. This interpretation is *usually* valid for more general algebraic approximants as will now be shown.

Let $\sum_{i=0}^{n} a_i(x) f(x)^i = O(x^{N+n})$.

It will be assumed that

(i) the $a_i(x)$ do not have a common factor (other than a constant)

(ii) $\sum_{i=0}^{n} a_i(x) f(x)^i$ is an irreducible polynomial in $(x, f(x))$.

(iii) $a_n(0) \neq 0$

(iv) $\frac{\partial}{\partial f}(\sum_{i=0}^{n} a_i(x) f(x)^i)|_{x=0} \neq 0$.

In the quadratic case these assumptions can be abandoned. This will be shown in Chapter 3.

Theorem 1.

(i) The equation

$$\sum_{i=0}^{n} a_i(x)\, y(x)^i = 0$$

defines a function $y(x)$ which, at the origin, has n distinct analytic branches.

(ii) The singularities of $y(x)$ (ignoring the point at infinity) come from two sources:

1. The roots of $a_n(x)$ which are infinitudes of one or more branches of $y(x)$.

2. The roots of the equation in x obtained by elimination of $y(x)$ from the simultaneous equations

$$\sum_{i=0}^{n} a_i(x)\, y(x)^i = 0$$
$$\frac{\partial}{\partial y}\left(\sum_{i=0}^{n} a_i(x)\, y(x)^i\right) = 0\,.$$

These points may be branch points of $y(x)$. $\sum_{i=0}^{n} a_i(x)y(x)^i = 0$ has multiple roots at and only at these points.

Proof. This theorem is a standard one in the theory of algebraic functions. The above statement was adapted from that of Hille ([6] Theorem 12.2.1) wherein can be found a detailed proof and discussion. See also Bliss [3]. □

At the origin $\sum_{i=0}^{n} a_i(x)y(x)^i$ has n roots, one of which is $f(0)$ (since $\sum_{i=0}^{n} a_i(x)f(x)^i|_0 = O(x^{N+n})|_0 = 0$) i.e. precisely one branch of $y(x)$ passes through $(0, f(0))$ and that branch is analytic in a neighbourhood of the origin. For the remainder of this chapter it will be understood that $y(x)$ means this unique branch of the multivalued function defined by $\sum_{i=0}^{n} a_i(x)y(x)^i = 0$.

Theorem 2. $y(x) = f(x) + O(x^{N+n})$
i.e. the power series expansions of $f(x)$ and $y(x)$ have the same first $N+n$ coefficients.

Proof.

Note that

$$\sum_{j=0}^{n} a_j(x)\, f(x)^j = O\left(x^{N+n}\right)$$
and
$$\sum_{j=0}^{n} a_j(x)\, y(x)^j = 0$$

so that

$$\frac{d^i}{dx^i}\left[\sum_j a_j(x)\, y(x)^j\right]\Big|_0 = 0 = \frac{d^i}{dx^i}\left[\sum_j a_j(x)\, f\left(x^j\right)\right]\Big|_0 \qquad i \in \{0, \ldots, N+n-1\}\,.$$

For $i = 1$

$$\left[\frac{\partial}{\partial y}\left(\sum a_j(x)\, y(x)^j\right)\frac{dy}{dx} + \left(\sum \frac{d}{dx}\left(a_j(x)\right) y(x)^j\right)\right]\Big|_0 = 0$$
$$\left[\frac{\partial}{\partial f}\left(\sum a_j(x)\, f(x)^j\right)\frac{df}{dx} + \left(\sum \frac{d}{dx}\left(a_j(x)\right) f(x)^j\right)\right]\Big|_0 = 0\,.$$

Differentiating again ($i = 2$) gives

$$\left[\frac{\partial}{\partial y}\left(\sum a_j(x)\,y(x)^j\right)\frac{d^2y}{dx^2}+\frac{d}{dx}\left(\frac{\partial}{\partial y}\left(\sum a_j(x)\,y(x)^j\right)\right)\frac{dy}{dx}\right.$$
$$\left.+\frac{d}{dx}\left(\sum\frac{d}{dx}(a_j(x))y(x)^j\right)\right]\Big|_0=0$$

$$\left[\frac{\partial}{\partial f}\left(\sum a_j(x)\,f(x)^j\right)\frac{d^2f}{dx^2}+\frac{d}{dx}\left(\frac{\partial}{\partial f}\left(\sum a_j(x)\,f(x)^j\right)\right)\frac{df}{dx}\right.$$
$$\left.+\frac{d}{dx}\left(\sum\frac{d}{dx}(a_j(x))\,f(x)^j\right)\right]\Big|_0=0\,.$$

In a general, more compact form we have

$$\left[\frac{\partial}{\partial y}\left(\sum a_j(x)\,y(x)^j\right)\frac{d^iy}{dx^i}+z_i\right]\Big|_0=0=\left[\frac{\partial}{\partial f}\left(\sum a_j(x)\,f(x)^j\right)\frac{d^if}{dx^i}+Z_i\right]\Big|_0 \tag{4}$$

$$i\in\{1,\ldots,N+n-1\}$$

where,

$$z_1=\sum\frac{d}{dx}(a_j(x))\,y(x)^j\,,$$

$$z_{i+1}=z_{i+1}\left(x,y(x),\frac{dy}{dx},\ldots,\frac{d^iy}{dx^i}\right)=\frac{dz_i}{dx}+\frac{d^iy}{dx^i}\frac{d}{dx}\left(\frac{\partial}{\partial y}\sum a_j(x)\,y(x)^j\right),$$

$$Z_{i+1}=z_{i+1}\left(x,f(x),\frac{df}{dx},\ldots,\frac{d^if}{dx^i}\right).$$

Since $y(0)=f(0)$ and $\frac{\partial}{\partial y}(\sum a_j(x)y(x)^j)|_0\neq 0$, equation (4) with $i=1$ gives $\left.\frac{df}{dx}\right|_0=\left.\frac{dy}{dx}\right|_0$, which with $i=2$ gives $\left.\frac{d^2f}{dx^2}\right|_0=\left.\frac{d^2y}{dx^2}\right|_0$. It follows that

$$\left.\frac{d^if}{dx^i}\right|_0=\left.\frac{d^iy}{dx^i}\right|_0\qquad i\in\{0,\ldots,N+n-1\}$$

i.e. $y(x)=f(x)+O(x^{N+n})$. □

So, if assumptions (i)—(iv) hold, then algebraic Hermite-Padé approximation generalises the idea behind Padé approximation i.e. finding a rational function with derivatives matching those of $f(x)$. Another theorem is now presented which, as well as giving information about behaviour when $f(x)$ is odd or even, indicates that assumptions (iii) and (iv) may often not be satisfied.

Theorem 3. $\forall A_0, \ldots, A_n \in \mathbf{N} \cup \{-1\}$.

1. If $f(x)$ is odd then there exists a non-trivial $(A_n, \ldots, A_0)$ Hermite-Padé form with either

(i)		or	(ii)	
$a_0(x)$	even		$a_0(x)$	odd
$a_1(x)$	odd		$a_1(x)$	even
$a_2(x)$	even		$a_2(x)$	odd
.			.	
.			.	
.			.	

2. If $f(x)$ is even then there exists a non-trivial $(A_n, \ldots, A_0)$ Hermite-Padé form with either

(i)		or	(ii)	
$a_0(x)$	even		$a_0(x)$	odd
$a_1(x)$	even		$a_1(x)$	odd
$a_2(x)$	even		$a_2(x)$	odd
.			.	
.			.	
.			.	

Proof.

1.
$$\sum a_i(x) f(x)^i = O\left(x^{N+n}\right) \tag{5}$$
$$\Rightarrow \sum a_i(-x) f(-x)^i = O\left((-x)^{N+n}\right) = O\left(x^{N+n}\right)$$
$$\Rightarrow \sum a_i(-x)(-1)^i f(x)^i = O\left(x^{N+n}\right) \tag{6}$$

Adding (5) and (6)
$$\sum \left(a_i(x) + (-1)^i a_i(-x)\right) f(x)^i = O\left(x^{N+n}\right).$$
So, either the Hermite-Padé form given by $\left\{a_i(x) + (-1)^i a_i(-x) : i \in \{0, \ldots, n\}\right\}$ is non-trivial (with $a_0(x)+a_0(-x)$ even, $a_1(x)-a_1(-x)$ odd and so on) or $a_0(x)$ is odd, $a_1(x)$ even etc.

2. Similar. □

One immediate consequence of Theorem 3 is that in the case where $f(x)$ is odd, n is even and there is only one linearly independent Hermite-Padé form of type $(A_n, \ldots, A_0)$ then either $a_n(0) = 0$ or $a_1(0) = 0$. Since $a_1(0) = 0$ and $f(0) = 0$ imply that $\frac{\partial}{\partial f}(\sum a_i(x)f(x)^i)|_{x=0} = 0$, in this case one of assumptions (iii) and (iv) is not satisfied.

5. Conclusion

Given the power series of a function $f(x)$ it has been shown that an algebraic Hermite-Padé form, if having certain properties, yields an approximation $y(x)$ (in the sense that $y(x) = f(x) + O(x^{N+n})$). There are, however, many outstanding problems, some of which we will attempt to answer (in the quadratic case at least) in the remainder of this thesis:

(i) Diagonalisation of the coefficient matrix in (3) requires $O(n^3)$ operations to generate the (n,n,n) form and so $O(n^4)$ operations to generate the sequence of forms $(0,0,0),(1,1,1),\ldots,(n,n,n)$. It would be a great advantage to develop an algorithm that would generate the (n,n,n) form from the $(n-1,n-1,n-1)$ form. Such an algorithm is presented in Chapter 2. It requires only $O(n^2)$ operations to generate the sequence $(0,0,0),\ldots,(n,n,n)$.

(ii) Theorem 2 which shows that $y(x) = f(x) + O(x^{N+n})$ is heavily dependent on the assumptions §4(i)—(iv). Unfortunately, in many examples these are not satisfied. In the quadratic case, however, they can be avoided. This is shown in Chapter 3.

(iii) What happens if there is more than one linearly independent Hermite-Padé form of type $(A_n, \ldots, A_0)$? Is there any discernable structure in the quadratic table? These questions are discussed in Chapter 6.

(iv) Do these approximations offer any advantage over the more usual Padé and Taylor approximations? This question will be discussed, with the aid of specific examples, in Chapters 4 and 5.

It is also of interest to ask whether any of the subsequent work can be generalised to cubic and other algebraic Padé-Hermite approximations or more importantly whether such a generalisation would be useful. Certainly Padé approximations have a distinct advantage over Taylor approximations in that they introduce poles as a possibility in the approximation and likewise quadratic approximations introduce branch point structure as a possibility. Cubic approximations, however, do not appear to add any new "structure" to the approximation of single-valued functions. This leads one to question whether the added complications in finding

and computing such approximations are worthwhile. Quadratic approximations to $\sqrt[3]{1+x}$ (which perform quite well (see Chapter 5, Example 3)) seem to support this comment.

6. References

1. G.A. Baker (1975): Essentials of Padé Approximants. New York: Academic Press.
2. G.A. Baker, D.S. Lubinsky (1987): *Convergence theorems for rows of differential and algebraic Hermite-Padé approximations.* J. Comput. Appl. Math., **18** : 29–52.
3. G. A. Bliss (1966): Algebraic Functions. New York: Dover Publications Inc.
4. P.B. Borwein (1987): *Quadratic and Higher Order Padé Approximants.* In: Colloquia Mathematics Societatis Janos Bolyai, **49** Alfred Haar Memorial Conference, Budapest (Hungary) 1985,: 213–224.
5. J. Della Dora, C. di Crescenzo (1984): *Approximations de Padé-Hermite.* Numer. Math. **43** : 23–57.
6. E. Hille (1962): Analytic Function Theory, vol II. Boston: Ginn and Company.
7. S.L. Loi, A.W. McInnes (1984): *An algorithim for the quadratic approximation.* J. Comput. Appl. Math. **11** :161–174.
8. S. Paszkowski (1987): *Recurrence Relations in Padé-Hermite Approximation.* J. Comput. Appl. Math. **19** : 99–107.

CHAPTER 2

A RECURRENCE ALGORITHM FOR QUADRATIC HERMITE-PADÉ FORMS

1. Introduction

A recurrence algorithm for calculating diagonal quadratic Hermite-Padé forms is derived. This algorithm has less restrictive conditions and applies to a larger class of functions than just those with normal systems as required by Paszkowski in [5], [6].

The Paszkowski algorithm is an algorithm for the more general algebraic Hermite-Padé form. It is a generalisation of the work of Padé [4], and the analysis of the algorithm is based on the Padé determinant [5]. Furthermore, the algorithm depends on the solution of a linear system of equations at each step.

Other recurrence relations have been derived by Della Dora and di Crescenzo [3], whose derivation is based on determinental identities, and by Borwein [1], [2], whose formulas are specific to the diagonal forms for $\exp(x)$ and $\log(x)$.

This algorithm is simply derived, is not complicated by the requirement to solve linear systems, and, since it requires less restrictive conditions applies to a much larger class of functions. The number of operations involved in each step is $O(n)$ (for the $[(n,n,n)]$ form), and hence the algorithm requires $O(n^2)$ operations to calculate all the diagonal forms up to the $[(n,n,n)]$ form. In §5, some examples are computed using the the algorithm. While this algorithm produces the correct results in symbolic computation, it should be noted that it appears to be unstable for numerical computation.

2. Notation

Let $f(x)$ be a given function with a known power series expansion about the origin.

(i) A quadratic Hermite-Padé form defined by

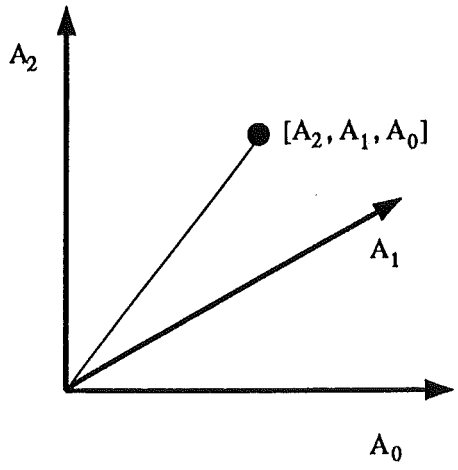
$$\sum_{i=0}^{2} a_i(x) f(x)^i = O(x^{A_0+A_1+A_2+2})$$

where $\deg(a_i(x)) \leq A_i$ will be expressed as $r(x,y) = a_2(x)y^2 + a_1(x)y + a_0(x)$ and often denoted by $[(A_2, A_1, A_0)]$.

(ii) The coefficient of $x^{A_0+A_1+A_2+2}$ in $\sum_{i=0}^{2} a_i(x) f(x)^i$ will be denoted by $E([(A_2, A_1, A_0)])$.

(iii) The coefficient of $x^i y^j$ in the polynomial $r(x,y)$ will be denoted by $\text{coeff}(r, x^i y^j)$.

(iv) To aid understanding we will imagine the quadratic Hermite-Padé forms to be elements of a 3–D lattice in the following arrangement.



It will be assumed throughout that the function $f(x)$ satisfies the following property.

Definition : Property A.

The function $f(x)$ satisfies :

(i) $E[(-1,0,0)] \neq 0$ for all such forms (where the only polynomial of degree -1 is the zero polynomial)

(ii) $E([(i,i,i)]), E([(i,i,i+1)]), E([(i,i+1,i+1)]) \neq 0$, for all possible such forms, $\forall i \in \mathbf{N}$.

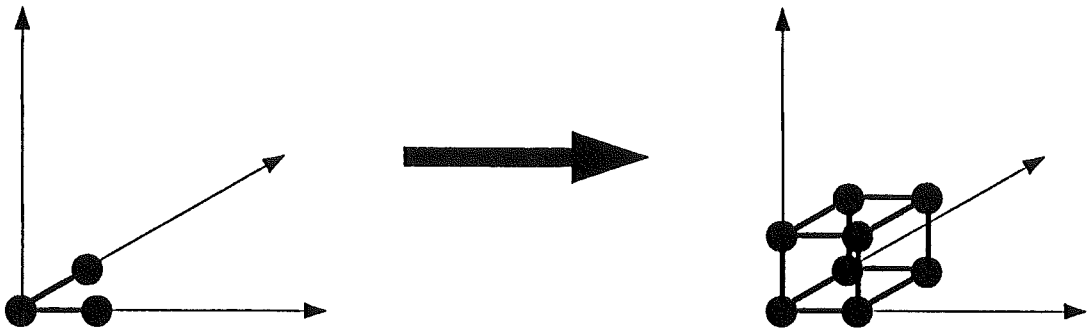
Note that normal systems (see Paszkowski [6]) satisfy the condition $E([(i,j,k)]) \neq 0, \forall i,j,k \in \mathbf{Z}^+$, and hence are included in the class of systems with Property A.

3. Discussion

This section will discuss the basis of the algorithm.

Suppose $[(0,0,0)], [(0,1,0)], [(0,0,1)]$ forms and their corresponding order coefficients, E, are known. Then forms of the type $[(i,j,k)]$ can be found $\forall i,j,k \in \{0,1\}$.

i.e. If these three forms are known then the remainder of the "cube" can be generated.



(i) A $[(1,0,0)]$ form.

Consider $r(x,y) = x[(0,0,0)] + \alpha_1[(0,1,0)] + \alpha_2[(0,0,1)]$ where :

α_1 is chosen so that coeff$(r, xy) = 0$

α_2 is chosen so that coeff$(r, x) = 0$.

Since $E([(0,0,0)]) \neq 0$ for all [(0,0,0)] forms the coefficient of xy in [(0,1,0)] and the coefficient of x in [(0,0,1)] are both non-zero, so such α_1, α_2 exist. This follows since if these coefficients were zero, then [(0,1,0)] and [(0,0,1)] would be [(0,0,0)] forms with order $O(x^3)$, which implies that $E([(0,0,0)]) = 0$ for these forms, contradicting Property A. Similarly, if the coefficient of y^2 in [(0,0,0)] were zero then this would be a [(-1,0,0)] form with order $O(x^2)$, again contradicting Property A. Hence $r(x,y) \not\equiv 0$ and since $r(x,y) = O(x^3)$, it follows that $r(x,y)$ is a $[(1,0,0)]$ form.

(ii) A $[(1,0,1)]$ form .

Set $r(x,y) = [(1,0,0)] + \alpha_1[(0,0,1)]$ where α_1 is chosen so that

$E([(1,0,0)]) + \alpha_1 E([(0,0,1)]) = 0$.

Since $E([(0,0,1)]) \neq 0$ such an α_1 exists and since $E([(0,0,0)]) \neq 0$ the coefficient of xy^2 in the [(1,0,0)] form is non-zero by the same argument as above, and so $r(x,y) \not\equiv 0$. It follows easily that $r(x,y)$ is a $[(1,0,1)]$ form.

(iii) A $[(0,1,1)]$ form .

Set $[(0,1,1)] = [(0,1,0)] + \alpha_1[(0,0,1)]$ where α_1 is chosen so that

$E([(0,1,0)]) + \alpha_1 E([(0,0,1)]) = 0$.

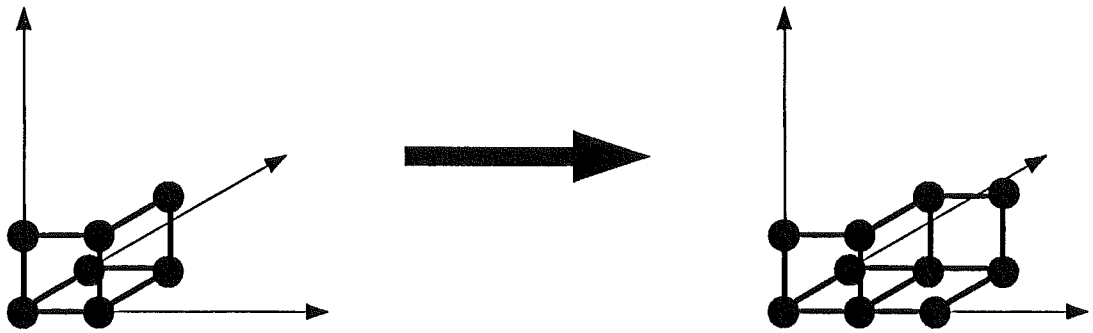
(iv) A $[(1,1,1)]$ form .

Set $[(1,1,1)] = [(1,0,1)] + \alpha_1[(0,1,1)]$ where α_1 is chosen so that

$E([(1,0,1)]) + \alpha_1 E([(0,1,1)]) = 0$.

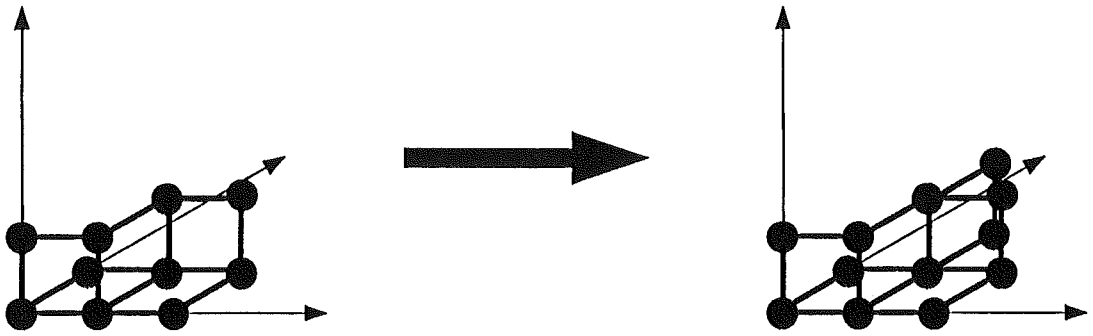
One can also generate a $[(1,1,0)]$ form in the same way although it will not be needed.

By symmetry it is clear that by using the same process the cube can be extended in the following way.



i.e. generate the forms $[(0,0,2)], [(1,1,2)], [(0,1,2)]$.

This can then be extended in the following way :



i.e. generate the forms $[(0,2,1)]$ and $[(1,2,1)]$.

It is now apparent that the initial pattern of three forms has been generated, but now centred at $(1,1,1)$ and that a simple recursion of this process will yield all the diagonal quadratic approximants. It is also clear that the number of operations involved at each step is only $O(n)$ so that it requires only $O(n^2)$ operations to generate the sequence $[(0,0,0)], [(1,1,1)], ..., [(n,n,n)]$.

4. Algorithm

Assume that the power series coefficients for $f(x)$ and $f(x)^2$ are given, and that the initial forms $[(0,0,0)], [(0,1,0)]$, and $[(0,0,1)]$ together with their corresponding order coefficients, E, have been independently calculated.

Step 1 : Set $[(0,1,1)] = [(0,1,0)] - \dfrac{E([(0,1,0)]}{E([(0,0,1)]}\,[(0,0,1)]$

$E([(0,1,1)]$ =coeff$([(0,1,1)], x^4)$

Set $n = 0$

Step 2 : Set $\mathbf{v} = (n,n,n), \mathbf{r} = (1,0,0), \mathbf{s} = (0,1,0), \mathbf{t} = (0,0,1)$

$z_1 = yx^n, z_2 = x^n, z_3 = x^{3n+3}$

Step 3 : Subroutine (see below)

Step 4 : Set $\mathbf{v} = (n,n,n+1), \mathbf{r} = (0,0,1), \mathbf{s} = (1,0,0), \mathbf{t} = (0,1,0)$

$z_1 = y^2x^n, z_2 = yx^n, z_3 = x^{3n+4}$

Step 5 : Subroutine

Step 6 : Set $\mathbf{v} = (n,n+1,n+1), \mathbf{r} = (0,1,0), \mathbf{s} = (0,0,1), \mathbf{t} = (1,0,0)$

$z_1 = x^{n+1}, z_2 = y^2x^n, z_3 = x^{3n+5}$

Step 7 : Subroutine

Step 8 : Set $n = n+1$

Go to Step 2

Subroutine

Step 1s : Set $[\mathbf{v}+\mathbf{r}] = x\,[\mathbf{v}] - \dfrac{\text{coeff}\,([\mathbf{v}], z_1)}{\text{coeff}\,([\mathbf{v}+\mathbf{s}], xz_1)}\,[\mathbf{v}+\mathbf{s}] - \dfrac{\text{coeff}\,([\mathbf{v}], z_2)}{\text{coeff}\,([\mathbf{v}+\mathbf{t}], xz_2)}\,[\mathbf{v}+\mathbf{t}]$

$E([\mathbf{v}+\mathbf{r}])$ =coeff$([\mathbf{v}+\mathbf{r}], z_3)$

Step 2s : Set $[\mathbf{v}+\mathbf{r}+\mathbf{t}] = [\mathbf{v}+\mathbf{r}] - \dfrac{E([\mathbf{v}+\mathbf{r}])}{E([\mathbf{v}+\mathbf{t}])}\,[\mathbf{v}+\mathbf{t}]$

$E([\mathbf{v}+\mathbf{r}+\mathbf{t}])$ =coeff$([\mathbf{v}+\mathbf{r}+\mathbf{t}], xz_3)$

Step 3s : Set $[\mathbf{v}+\mathbf{r}+\mathbf{s}+\mathbf{t}] = [\mathbf{v}+\mathbf{r}+\mathbf{t}] - \dfrac{E([\mathbf{v}+\mathbf{r}+\mathbf{t}])}{E([\mathbf{v}+\mathbf{s}+\mathbf{t}])}\,[\mathbf{v}+\mathbf{s}+\mathbf{t}]$

$E([\mathbf{v}+\mathbf{r}+\mathbf{s}+\mathbf{t}])$ =coeff$([\mathbf{v}+\mathbf{r}+\mathbf{s}+\mathbf{t}], x^2z_3)$

Note that coeff($[\mathbf{v}+\mathbf{s}], xz_1$) $\neq 0$, since if it were zero then $[\mathbf{v}+\mathbf{s}]$ would be a $[\mathbf{v}]$ form with $E([\mathbf{v}]) = 0$, contradicting Property A. Similarly coeff($[\mathbf{v}+\mathbf{t}]$) $\neq 0$, and hence the conditions of Property A are sufficient for the operation of this algorithm.

5. Examples

The algorithm has been used, coded in MACSYMA, on the functions $\exp(x), \log(1+x), \sin(x)$ and has been found to work satisfactorily. Following are the calculations for Steps 1–3 in some detail for $f(x) = \exp(-x)$ (note that the forms have been normalised at each step to give integer coefficients).

Given

$[(0,0,0)] = y^2 - 2y + 1$

$[(0,1,0)] = y^2 + 2xy - 1$

$[(0,0,1)] = y^2 - 4y - 2x + 3$

and

E([(0,1,0)]) = -1/3

E([(0,0,1)]) = -2/3

Then

Step 1

$[(0,1,1)] = 2[y^2 + 2xy - 1 - \frac{1/3}{2/3}(y^2 - 4y - 2x + 3)] = y^2 + (4x+4)y + 2x - 5$

$E([(0,1,1)] = 1/6$

Step 1s

$[(1,0,0)] = 2[x(y^2-2y+1)+(y^2+2xy-1)+(1/2)(y^2-4y-2x+3)] = (2x+3)y^2 - 4y + 1$

$E([(1,0,0)]) = 2/3$

Step 2s

$[(1,0,1)] = (1/2)[(2x+3)y^2 - 4y + 1 + \frac{2/3}{2/3}(y^2 - 4y - 2x + 3)] = (x+2)y^2 - 4y - x + 2$

$E([(1,0,1)]) = -1/6$

Step 3s

$$[(1,1,1)] = (x+2)y^2 - 4y - x + 2 + \frac{1/6}{1/6}\,(y^2 + (4x+4)y + 2x - 5) = (x+3)y^2 + 4xy + x - 3$$

$$E([(1,1,1)] = 1/30$$

When the calculations are performed numerically rather than symbolically however, a rapid build up of calculation error is evident. Several diagonal forms for $\log(1+x)$ are given here, calculated both exactly using MACSYMA and numerically using PC-Matlab with 16 digit precision. Although all the values are available to at least 16 digits only enough precision is given to indicate the numerical error. Note also that the normalisation $\|[(n,n,n)]\|_2 = 1$ is used (where $\|[(n,n,n)]\|_2$ means the ℓ_2 norm of the vector formed from the coefficients of $[(n,n,n)]$).

The $[(2,2,2)]$ form

Coefficient	PC-Matlab	MACSYMA
y^2	-1.84812331090 $\times 10^{-1}$	1.84812331090 $\times 10^{-1}$
xy^2	-1.84812331090 $\times 10^{-1}$	-1.84812331090 $\times 10^{-1}$
x^2y^2	3.08020551818 $\times 10^{-2}$	3.08020551816 $\times 10^{-2}$
y	2.90 $\times 10^{-12}$	0
xy	-5.54436993269 $\times 10^{-1}$	-5.54436993270 $\times 10^{-1}$
x^2y	-2.77218496635 $\times 10^{-1}$	-2.77218496635 $\times 10^{-1}$
1	0	0
x	-2.90 $\times 10^{-12}$	0
x^2	7.39249324361 $\times 10^{-1}$	7.39249324360 $\times 10^{-1}$

The $[(4,4,4)]$ form

Coefficient	PC-Matlab	MACSYMA
y^2	6.10111 $\times 10^{-2}$	6.10109 $\times 10^{-2}$
xy^2	1.22022 $\times 10^{-1}$	1.22022 $\times 10^{-1}$
x^2y^2	2.03374 $\times 10^{-2}$	2.03369 $\times 10^{-2}$
x^3y^2	-4.06739 $\times 10^{-2}$	-4.06739 $\times 10^{-2}$
x^4y^2	6.77897 $\times^{-4}$	6.77898 $\times 10^{-4}$
y	-1.56 $\times 10^{-6}$	0
xy	3.99958 $\times 10^{-1}$	3.99960 $\times 10^{-1}$
x^2y	5.99940 $\times 10^{-1}$	5.99940 $\times 10^{-1}$
x^2y	1.83033 $\times 10^{-1}$	1.83033 $\times 10^{-1}$
x^4y	-8.47371 $\times 10^{-3}$	-8.47373 $\times 10^{-3}$
1	0	0
x	1.56 $\times 10^{-6}$	0
x^2	-4.60970 $\times 10^{-1}$	-4.60971 $\times 10^{-1}$
x^3	-4.60972 $\times 10^{-1}$	-4.60971 $\times 10^{-1}$
x^4	2.93756 $\times 10^{-2}$	2.93756 $\times 10^{-2}$

The $[(5,5,5)]$ form

Coefficient	PC-Matlab	MACSYMA
y^2	3.76 $\times 10^{-5}$	0
xy^2	6.672 $\times 10^{-2}$	6.660 $\times 10^{-2}$
x^2y^2	1.333 $\times 10^{-1}$	1.332 $\times 10^{-1}$
x^3y^2	5.392 $\times 10^{-2}$	5.391 $\times 10^{-2}$
x^4y^2	-1.270 $\times 10^{-2}$	-1.269 $\times 10^{-2}$
x^5y^2	1.059 $\times 10^{-4}$	1.057 $\times 10^{-4}$
y	-2.541 $\times 10^{-1}$	-2.540 $\times 10^{-1}$
xy	-6.349 $\times 10^{-1}$	-6.349 $\times 10^{-1}$
x^2y	-3.445 $\times 10^{-1}$	-3.448 $\times 10^{-1}$
x^3y	1.180 $\times 10^{-1}$	1.177 $\times 10^{-1}$
x^4y	7.874 $\times 10^{-2}$	7.866 $\times 10^{-2}$
x^5y	-1.451 $\times 10^{-3}$	-1.448 $\times 10^{-3}$
1	0	0
x	2.540 $\times 10^{-1}$	2.540 $\times 10^{-1}$
x^2	5.079 $\times 10^{-1}$	5.080 $\times 10^{-1}$
x^3	4.503 $\times 10^{-2}$	4.544 $\times 10^{-2}$
x^4	-2.087 $\times 10^{-1}$	-2.085 $\times 10^{-1}$
x^5	5.435 $\times 10^{-3}$	5.424 $\times 10^{-3}$

6. Conclusion

Numerical problems aside, this algorithm gives a recursive procedure for calculating quadratic Hermite-Padé forms for many functions. The algorithm operates successfully for functions whose systems are not necessarily normal, such as $\log(1+x)$ and $\sin(x)$.

Although a procedure for calculating diagonal forms is given, it is clear that the algorithm could be easily generalised to achieve any [(l,m,n)] form by constructing the cubes in an appropriate manner to reach this point in the lattice. Similar concepts are mentioned by Padé [4], and Paszkowski [5], [6]. Property A would have to be modified in this case, to take account of the path through the lattice, but it is clear that a full normal system would not be required unless every point in the lattice was to be computed.

It is also clear that these principles could easily be extended to the more general algebraic Hermite-Padé forms studied by Paszkowski [6].

It is of interest to note that the relationships obtained in the algorithm generalise in a natural way some of the Frobenius identities familiar in Padé approximation, as was noted by Paszkowski [6].

7. References

1. P.B. Borwein, (1986) : *Quadratic Hermite-Padé Approximation to the Exponential Function.* Constr. Approx. **2**:291–302.
2. P.B. Borwein, (1987) : *Quadratic and Higher Order Padé Approximants.* In: Colloquia Mathematica Societatis Janos Bolyai, **49** Alfred Haar Memorial Conference, Budapest (Hungary) 1985, : 213–224.
3. J. Della Dora, C. di Crescenzo, (1984) : *Approximants de Padé-Hermite.* Numer. Math. **43**:23–57.
4. H. Padé, (1894) : *Sur la généralisation des fractions continues algebriques.* J. Pures et Appl. **10**:291–329.
5. S. Paszkowski, (1982) : *Quelques Algorithms de l'Approximation de Padé-Hermite.* Publ. ANO-89, Univ. Lille I.
6. S. Paszkowski, (1987) : *Recurrence Relations in Padé-Hermite Approximation.* J. Comput. Appl. Math. **19**:99–107.

CHAPTER 3

THE EXISTENCE AND LOCAL BEHAVIOUR OF THE QUADRATIC HERMITE-PADÉ APPROXIMATION

1. Introduction

In the case of the (A_1, A_0) Padé approximation (Baker [1]) it is well known that if $a_1(0) \neq 0$ then $y(x) = -a_0(x)/a_1(x)$ satisfies $y(x) = f(x) + O(x^{A_0+A_1+1})$. However, in the quadratic case it is not obvious that $a_2(x)y(x)^2 + a_1(x)y(x) + a_0(x) = 0$ yields even an analytic approximation to $f(x)$, still less that it defines a function $y(x)$ such that $y(x) = f(x) + O(x^{A_0+A_1+A_2+2})$. The purpose of this chapter is to show that an analogue of the Padé result is in fact true.

2. Notation

It will be assumed that

$$\sum_{j=0}^{2} a_j(x) f(x)^j = O\left(x^{N+2}\right) \tag{1}$$

where $N \geq \sum_j A_j$ and that $\sum_j |a_j(0)| \neq 0$.

The following notation will be used:

(i) An approximation derived from $\sum_{j=0}^{2} a_j(x)f(x)^j = O(x^{N+2})$ will be referred to as a (A_2, A_1, A_0) (quadratic) approximation to $f(x)$.

(ii) By $\sqrt{D(x)}$ we mean the principal square root of $D(x)$.

(iii) Let $D(x) = a_1(x)^2 - 4a_2(x)a_0(x)$. If $\sum_j a_j(x)y(x)^j = 0$ then

$$y(x) = \frac{-a_1(x) \pm \sqrt{D(x)}}{2a_2(x)} \quad \text{and}$$

$$\pm\sqrt{D(x)} = 2a_2(x)y(x) + a_1(x) = \frac{\partial}{\partial y}\left(\sum_j a_j(x)y(x)^j\right).$$

3. The Principal Results:

The problem divides itself into two cases, the case $D(0) \neq 0$ and the case $D(0) = 0$.

3.1 The case $D(0) \neq 0$.

Theorem 1. If $D(0) \neq 0$ then there exists a unique function $y(x)$, analytic in a neighbourhood of the origin, satisfying $\sum_j a_j(x)y(x)^j = 0$ and $y(0) = f(0)$.

Proof. The existence of a function $y(x)$, analytic about the origin, satisfying $\sum_j a_j(x)y(x)^j = 0$ follows from standard algebraic function theory (see for example Hille [4], Theorem 12.2.1). However in this special case it is easier to argue as follows.

Suppose $a_2(0) \neq 0$. The two possible expressions for $y(x)$ in a neighbourhood of the origin are given by

$$y(x) = \frac{-a_1(x) - \sqrt{D(x)}}{2a_2(x)}$$

or

$$y(x) = \frac{-a_1(x) + \sqrt{D(x)}}{2a_2(x)}.$$

Since $D(0) \neq 0$ these are both analytic in a neighbourhood of the origin. Exactly one of them satisfies $y(0) = f(0)$ because $a_2(0)f(0)^2 + a_1(0)f(0) + a_0(0) = 0$

$$\Rightarrow f(0) = \frac{-a_1(0) \pm \sqrt{D(0)}}{2a_2(0)}.$$

Suppose $a_2(0) = 0$. Then $a_1(0) \neq 0$ (again since $D(0) \neq 0$). Near the origin the two possible expressions for $y(x)$ can be written as :

$$y(x) = \frac{-a_1(x) + a_1(x)\sqrt{1 - \dfrac{4a_2(x)a_0(x)}{a_1(x)^2}}}{2a_2(x)} \tag{2}$$

or

$$y(x) = \frac{-a_1(x) - a_1(x)\sqrt{1 - \dfrac{4a_2(x)a_0(x)}{a_1(x)^2}}}{2a_2(x)}. \tag{3}$$

The right hand side of (3) is unbounded as $x \to 0$ so we can exclude this possibility. Since $a_2(0) = 0$, close to the origin we can apply the binomial theorem to get from (2) the convergent power series (analytic in a neighbourhood of the origin) expression for $y(x)$:

$$y(x) = \left(-a_1(x) + a_1(x)\left(1 + \sum_{i=1}^{\infty} \mu_i \left(\frac{4a_2(x)a_0(x)}{a_1(x)^2}\right)^i\right)\right)/2a_2(x)$$

$$= \sum_{i=1}^{\infty} \mu_i \left(\frac{4a_2(x)a_0(x)}{a_1(x)^2}\right)^{i-1} \frac{2a_0(x)}{a_1(x)}.$$

Noting that $\mu_1 = -\frac{1}{2}$ it follows that

$$y(x) = \begin{cases} \dfrac{-a_1(x) + a_1(x)\sqrt{1 - \dfrac{4a_2(x)a_0(x)}{a_1(x)^2}}}{2a_2(x)} & x \neq 0 \\ -\dfrac{a_0(x)}{a_1(x)} & x = 0 \end{cases}.$$

is the only function, analytic in a neighbourhood of the origin, satisfying
$\sum_j a_j(x)y(x)^j = 0 \quad \text{with} \quad y(0) = f(0)$. □

Theorem 2. If $D(0) \neq 0$ then there exists a unique function $y(x)$, analytic in a neighbourhood of the origin, satisfying $\sum_j a_j(x)y(x)^j = 0$ such that

$$y(x) = f(x) + O(x^{N+2}) .$$

Proof. This proof is identical to that of Chapter 1 Theorem 2. Note that Theorem 1 above which guarantees the exist of a unique analyic y(x) is dependent only on the assumptions $\sum_j |a_j(0)| \neq 0$ and $D(0) \neq 0$. □

3.2 The case $D(0) = 0$.

We now investigate the case $D(0) = 0$. This implies that $a_2(0) \neq 0$ (since if $D(0) = a_1(0)^2 - 4a_2(0)a_0(0) = 0$ and $a_2(0) = 0$ then $a_1(0) = 0$, which with $a_2(0)f(0)^2 + a_1(0)f(0) + a_0(0) = 0$ gives $a_0(0) = 0$. This contradicts the assumption that $\sum_j |a_j(0)| \neq 0$).

Bearing in mind that $y(x) = \dfrac{-a_1(x) \pm \sqrt{D(x)}}{2a_2(x)}$ and $\sqrt{D(x)}$ is not now analytic at the origin this case does not seem well-behaved, but such is not the case. Certainly if $D(x)$ has a root of odd multiplicity $2r + 1$, say, at the origin then any $y(x)$ satisfying $\sum_j a_j(x)y(x)^j = 0$ is not analytic at the origin since

$$\lim_{t\to 0} \frac{d^{r+1}}{dx^{r+1}} \frac{\sqrt{xg(x)}}{a_2(x)} x^r \Big|_t \to \infty \quad (g(0) \neq 0) .$$

[Take for example $x\sqrt{x}$. Then $\dfrac{d^2}{dx^2}(x\sqrt{x}) = \dfrac{3}{4\sqrt{x}}$. This generalises easily (using the Leibnitz rule) to the above]. However, this case never occurs in practice.

Firstly, it is necessary to treat two special cases:

(i) Suppose $a_0(x) \equiv 0$.

Then

$$a_2(x) f(x)^2 + a_1(x) f(x) = O\left(x^{N+2}\right)$$
$$\Rightarrow (a_2(x) f(x) + a_1(x)) f(x) = O\left(x^{N+2}\right)$$

so that

$$\left.\begin{array}{ll} -a_1(x)/a_2(x) & = f(x)+O\left(x^R\right) \\ 0 & = f(x)+O\left(s^S\right) \end{array}\right\} \quad \text{where} \quad R+S=N+2.$$

Choosing

$$\begin{cases} y(x) = -\dfrac{a_1(x)}{a_2(x)} & \text{if } R>S \\ y(x) = 0 & \text{otherwise} \end{cases}$$

gives $y(x)$ such that

$$\sum_i a_i(x)\,y(x)^i = 0$$

and $y(x) = f(x) + O(x^{\max\{R,S\}})$.

(Clearly $\max\{R,S\} \geq \dfrac{N}{2}+1$).

(ii) Suppose $D(x) \equiv 0$. Then

$$\begin{aligned} & a_2(x)\,f(x)^2 + a_1(x)\,f(x) + a_0(x) = O\left(x^{N+2}\right) \\ \Rightarrow & (2a_2(x)\,f(x) + a_1(x))^2 = 4a_2(x)\,O\left(x^{N+2}\right) = O\left(x^{N+2}\right) \\ \Rightarrow & y(x) = -\frac{a_1(x)}{2a_2(x)} = f(x) + O\left(x^T\right), T = \min\left\{\mathrm{t} \in \mathbf{N} : \mathrm{t} \geq \frac{\mathrm{N}}{2}+1\right\} \end{aligned}$$

and $\sum_i a_i(x)y(x)^i = 0$.

It will be assumed for the remainder of this section that neither $D(x) \equiv 0$ nor $a_0(x) \equiv 0$.

Theorem 3. Let $C_i = \deg(a_i(x))$. Then $D(x)$ never has a root of multiplicity greater than $\sum_i C_i$ at the origin.

Proof. Let $\sum_i C_i = M$ and suppose $D(x) = x^{M+1}p_r(x)$, $p_r(x)$ a polynomial of degree r. Since $a_2(x), a_0(x) \not\equiv 0$ then

$$a_1(x)^2 = x^{M+1}p_r(x) + q_s(x) \tag{4}$$

where $q_s(x)$ is a (nonzero) polynomial of degree s. We must have $M+1+r = 2C_1$ (since $C_2 + C_0 \leq M < M+1$) and $q_s(x) = 4a_0(x)a_2(x)$.

Also $s + C_1 = C_0 + C_2 + C_1 < M + 1 = 2C_1 - r \Rightarrow s + r < C_1$.

Differentiating (4)

$$2a_1(x)a_1'(x) = x^M((M+1)p_r(x) + xp_r'(x)) + q_s'(x)$$

$$\begin{aligned} \Rightarrow 2xa_1(x)\,a_1'(x) &= x^{M+1}\left((M+1)\,p_r(x) + xp_r'(x)\right) + xq_s'(x) \\ &= x^{M+1}\bar{p}_r(x) + \bar{q}_s(x) \end{aligned} \tag{5}$$

where

$$\bar{p}_r(x) \quad = (M+1)\,p_r(x) + xp_r'(x) \quad \text{(degree } r\text{)}$$

$$\bar{q}_s(x) \quad = xq_s'(x) \qquad\qquad\qquad \text{(degree } s\text{)}\,.$$

From (4) and (5) (eliminating the term in x^{M+1})

$$a_1(x)\left(\bar{p}_r(x)\,a_1(x) - 2p_r(x)\,xa_1'(x)\right) = \begin{vmatrix} q_s(x) & p_r(x) \\ \bar{q}_s(x) & \bar{p}_r(x) \end{vmatrix}. \tag{6}$$

The left-hand side of (6) either has degree $\geq C_1$ or is identically zero, while the right-hand side has degree $\leq s + r < C_1$ .It follows that

$$\bar{p}_r(x)\,a_1(x) - 2p_r(x)\,xa_1'(x) = 0 = q_s(x)\,\bar{p}_r(x) - p_r(x)\,\bar{q}_s(x)\,.$$

Hence

$$\frac{a_1'(x)}{a_1(x)} = \frac{\bar{p}_r(x)}{2xp_r(x)} = \frac{\bar{q}_s(x)}{2xq_s(x)} = \frac{q_s'(x)}{2q_s(x)}$$

and integrating gives

$$a_1(x) = k\sqrt{q_s(x)} \qquad , k \in \mathbf{R}.$$

But $\deg\sqrt{q_s(x)} = s/2 < C_1$ so the result is proved by contradiction. □

Theorem 4. $D(x)$ never has a root of odd multiplicity at the origin.

Proof. Suppose $D(x) = x^{2s+1}g(x), g(0) \neq 0$. By Theorem 3 it can be assumed that $2s + 1 < N + 1$. Then

$$\frac{d^{2s+1}}{dx^{2s+1}}\,D(x)\,|_0 \neq 0\,. \tag{7}$$

$$\frac{d^i}{dx^i}D(x)\,|_0 = 0 \qquad i \in \{0,\ldots,2s\}\,. \tag{8}$$

Let $G(x) = (2a_2(x)f(x) + a_1(x))$.

Then

$$a_2(x)\,f(x)^2 + a_1(x)\,f(x) + a_0(x) = O\left(x^{N+2}\right)$$

$$\Rightarrow G(x)^2 - D(x) = 4a_2(x)\,O\left(x^{N+2}\right) = O\left(x^{N+2}\right). \tag{9}$$

From (8) and (9)

$$\frac{d^i}{dx^i} G(x)^2 \,|_0 \;=\; 0 \qquad i \in \{0,\ldots,2s\} \tag{10}$$

$$\Rightarrow \sum_{j=0}^{i} \binom{i}{j} \frac{d^j}{dx^j} G(x)\, \frac{d^{i-j}}{dx^{i-j}} G(x)\,|_0 = 0 \qquad i \in \{0,\ldots,2s\}$$

$$\Rightarrow \frac{d^i}{dx^i} G(x)\,|_0 = 0 \qquad i \in \{0,\ldots,s\}.$$

[Expanding the first few equations:

$$\begin{aligned}
G(x)^2\,|_0 \;&=\; 0 \Rightarrow G(x)\,|_0 \;=\; 0 \\
\frac{d}{dx} G(x)^2 \,\Big|_0 \;&=\; 0 \Rightarrow \left[G(x)\,G'(x) + G'(x)\,G(x)\right] \Big|_0 \;=\; 0 \\
\frac{d^2}{dx^2} G(x)^2 \,\Big|_0 \;&=\; 0 \Rightarrow \left[G(x)\,G''(x) + 2G'(x)^2 + G''(x)\,G(x)\right] \Big|_0 \;=\; 0 \\
&\qquad \Rightarrow G'(x)\,|_0 \;=\; 0\Big]
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow \frac{d^{2s+1}}{dx^{2s+1}} G(x)^2\,|_0 \;=\; 0 \\
&\Rightarrow \frac{d^{2s+1}}{dx^{2s+1}} D(x)\,|_0 \;=\; 0.
\end{aligned}$$

Hence the result is shown by contradiction. □

Theorem 5. If $D(x) = x^{2s}g(x), g(0) \neq 0, 2s < N+1$ then either

$$y(x) \;=\; \frac{-a_1(x) + x^s\sqrt{g(x)}}{2a_2(x)}$$

or

$$y(x) \;=\; \frac{-a_1(x) - x^s\sqrt{g(x)}}{2a_2(x)}$$

satisfies

$\sum a_j(x)y(x)^j = 0$ and $y(x) = f(x) + O(x^{N+2-s})$.

Proof. Let $h(x) = x^s\sqrt{g(x)}$.

Then $G(x)^2 - h(x)^2 = O(x^{N+2})$ (G(x) as defined in the proof of Theorem 4)

$$\Rightarrow \quad \frac{d^i}{dx^i}G(x)^2\Big|_0 = \frac{d^i}{dx^i}h(x)^2\Big|_0, \quad i \in \{0,\ldots,N+1\}\,. \tag{11}$$

Also $\frac{d^i}{dx^i}\,G(x)^2|_0 = 0 = \frac{d^i}{dx^i}\,h(x)^2|_0 \quad i \in \{0,\ldots,2s-1\}$

so $\frac{d^i}{dx^i}\,G(x)|_0 = 0 = \frac{d^i}{dx^i}\,h(x)|_0 \quad i \in \{0,\ldots,s-1\}$

(using ideas in the proof of Theorem 4).

But $\frac{d^{2s}}{dx^{2s}}\,G(x)^2|_0 = \frac{d^{2s}}{dx^{2s}}\,h(x)^2|_0 \neq 0$

$$\Rightarrow \left(\frac{d^s}{dx^s}G(x)\Big|_0\right)^2 = \left(\frac{d^s}{dx^s}h(x)\Big|_0\right)^2 \neq 0\,.$$

Now choose $t(x) = h(x)$ or $t(x) = -h(x)$ so that $\frac{d^s}{dx^s}\,G(x)|_0 = \frac{d^s}{dx^s}\,t(x)|_0$.

Then equation (11) with $i = 2s+1$ gives

$$\begin{aligned}&\frac{d^{2s+1}}{dx^{2s+1}}G(x)^2\Big|_0 = \frac{d^{2s+1}}{dx^{2s+1}}t(x)^2\Big|_0\\ \Rightarrow &\left(\frac{d^s}{dx^s}G(x)\,\frac{d^{s+1}}{dx^{s+1}}G(x)\right)\Big|_0 = \left(\frac{d^s}{dx^s}t(x)\,\frac{d^{s+1}}{dx^{s+1}}t(x)\right)\Big|_0\,.\end{aligned}$$

We progress in this way up to $i = N+1$ (Note that if $2s = N+1$ this procedure is not required).

It follows that

$$\frac{d^i}{dx^i}G(x)\Big|_0 = \frac{d^i}{dx^i}t(x)\Big|_0, \quad i \in \{0,\ldots,N+1-s\}$$

i.e. $2a_2(x)f(x) + a_1(x) = t(x) + O(x^{N+2-s})$.

Since $a_2(0) \neq 0$, defining

$$y(x) = -\frac{a_1(x) - t(x)}{2a_2(x)}$$

gives $y(x) = f(x) + O(x^{N+2-s})$. □

4. Illustrative Examples

The seemingly exceptional cases covered by the previous theorems do frequently occur as is shown below.

Example 1. Let $f(x) = \log(1+x)$.
Then $xf(x)^2 + (-6x-12)f(x) + 12x = O(x^5)$.
Also

$$y(x) = \frac{6x+12-\sqrt{(6x+12)^2-48x^2}}{2x}$$
$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{29x^5}{144} + \dots$$

i.e. $y(x) = f(x) + O(x^5)$ (cf. the case $a_2(0) = 0$ in the proof of Theorem 1).

Example 2. Let $f(x) = \log(1+x)$.
Then $(x^2-6x-6)f(x)^2 + (-9x^2-18x)f(x) + 24x^2 = O(x^8)$.
Note that $D(x) = -15x^4 + 900x^3 + 900x^2$.

Also

$$y(x) = \frac{9x^2+18x-x\sqrt{-15x^2+900x+900}}{2(x^2-6x-6)}$$
$$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5} - \frac{x^6}{6} + \frac{1543x^7}{10800} + \dots$$

i.e. $y(x) = f(x) + O(x^7)$ as predicted by Theorem 5.

Example 3. Let $f(x) = 2x^5 + \frac{x^6}{2} - \frac{x^7}{8} + \frac{x^8}{16} - \frac{5x^9}{128} + \frac{7x^{10}}{256} - \frac{21x^{11}}{1024} + \frac{33x^{12}}{2048} + O(x^{13})$.

Then $f(x)^2 - 2x^5 f(x) - x^{11} = O(x^{18})$.
Note that $D(x) = 4x^{11} + 4x^{10}$.
Also

$$y(x) = x^5 + x^5\sqrt{x+1}$$

$$= 2x^5 + \frac{x^6}{2} - \frac{x^7}{8} + \frac{x^8}{16} - \frac{5x^9}{128} + \frac{7x^{10}}{256}$$

$$- \frac{21x^{11}}{1024} + \frac{33x^{12}}{2048} - \frac{429x^{13}}{32768} + \dots$$

i.e. $y(x) = f(x) + O(x^{13})$ as predicted by Theorem 5.

5. Conclusion

These results show that given
$a_2(x)f(x)^2 + a_1(x)f(x) + a_0(x) = O(x^{N+2})$, $\sum_{i=0}^{2} |a_i(0)| \neq 0$ then we can always find $y(x)$ such that $\sum_{i=0}^{2} a_i(x)y(x)^i = 0$ and $y(x) = f(x) + O(x^K)$ where K is, at worst $\frac{N}{2} + 1$ .This is summarised in the following table.

	K
$D(0) \neq 0$	$N+2$
$D(x) = x^{2s}g(x)$ where $g(0) \neq 0$, $2s < N+1$, $a_0(x) \not\equiv 0$	$N+2-s$
$D(0) = 0$ and $a_0(x) \equiv 0$	$\min\{k \in \mathbf{N} : k \geq \frac{N}{2} + 1\}$
$D(x) \equiv 0$	$\min\{k \in \mathbf{N} : k \geq \frac{N}{2} + 1\}$

Table 1

It is hoped that this work will be useful in attempting to extend convergence results such as that given by Baker and Lubinsky [2] to the so-called "non-normal" case in quadratic approximation.

Acknowledgement
The author wishes to thank Prof. R.P. Kerr for the proof of Theorem 3.

6. References

1. G.A. Baker, (1975) : Essentials of Padé Approximants. New York: Academic Press.
2. G.A. Baker, D.S. Lubinsky (1987) : *Convergence theorems for rows of differential and algebraic Hermite-Padé approximations.* J. Comput. Appl. Math., **18** : 29-52.
3. J. Della Dora, C. di Crescenzo (1984) : *Approximants de Padé-Hermite.* Numer. Math., **43** : 23-57.
4. E. Hille (1962) : Analytic Function Theory, vol II. Boston : Ginn and Company.

CHAPTER 4

SOME QUALITATIVE RESULTS FOR THE QUADRATIC HERMITE-PADÉ APPROXIMATION

1. Introduction

In the previous chapter the existence of local quadratic approximations to any analytic function $f(x)$ was established. The questions which we now attempt to answer, mainly by means of illustrative examples, are these:

(i) Can the approximation $y(x)$, unique in some neighbourhood of the origin, be extended to approximate $f(x)$ over some larger region?

(ii) Does this method of approximation give significantly better results than more traditional methods (eg. Padé and Taylor approximations) and if so, for what types of functions?

2. Discussion

In what follows it will be assumed that $\sum_{i=0}^{2} a_i(x)y(x)^i = 0$ and that

(i) the $a_i(x)$ do not have a common factor

(ii) $\sum a_i(x)y(x)^i$ cannot be factorised.
i.e. $\nexists$ polynomials $p(x), q(x), r(x), s(x)$, such that

$$\sum a_i(x)y(x)^i \equiv (p(x)y(x) + q(x))(r(x)y(x) + s(x)) \ .$$

These are not serious restrictions since:

(1) If the $a_i(x)$ have a (maximal) common factor $x^r p(x), p(0) \neq 0$ and $a_i(x) = x^r p(x) g_i(x)$ then $\sum_i g_i(x) f(x)^i = O(x^{N+2-r})$.

(2) If $\sum_i a_i(x) f(x)^i$ can be factorised then the approximation $y(x)$ is rational (and the theorems (see Chapter 3) regarding order of accuracy still apply).

It then follows (Hille [3] Theorem 12.2.1) that $y(x)$ (as defined in Chapter 3) is analytic everywhere except possibly at the points $x \in \mathbf{C}$ such that $a_2(x) = 0$ (poles) and the points $x \in \mathbf{C}$ such that $D(x) = 0$ (branch points).

Theorem 1. $y(x)$ is single valued and analytic in any simply connected neighbourhood of the origin not containing any of the above points.

Proof. The result follows easily from a standard complex variable result, namely that a set B is simply connected if and only if every function analytic and with no zeroes in B has an analytic square root in B (see for example Curtis [2] Theorem 12.8.1). □

It will be assumed that $f(x)$, the function being approximated is single-valued, or at least, that we wish to approximate only one of its Riemann sheets. It is then clearly necessary to restrict

the region of approximation, R, so that $y(x)$ is single valued on R. This consideration, with some additional information about $f(x)$ (for instance, the knowledge that $f(x)$ is analytic on some region, or the approximate location of any singularities) turns out to give enough information to accurately approximate some functions over a wide area. It is important at this point to realise that this "additional information" is necessary simply to ensure that the approximation problem is well-defined. The function $f(x)$ is not completely defined by it's power series at the origin, it may have many possible analytic continuations and so more information is needed to decide which possible continuation of the approximation is likely to be closest to $f(x)$.

3. Examples

3.1 Example 1

The (2,2,2) approximation to log(1+x) Consider the $(2,2,2)$ quadratic approximation to the principal branch of $f(x) = \log(1+x)$ (with a cut taken along $\{x \in \mathbf{R} : x \in (-\infty, -1]\}$). Note that:

(i) $(x^2-6x-6)f(x)^2-(9x^2+18x)f(x)+24x^2 = O(x^8)$ so that (using results from Chapter 3) the approximation is

$$y(x) = \frac{9x^2+18x-x\sqrt{-15x^2+900x+900}}{2(x^2-6x-6)}$$

$$\text{with} \quad y(x) = f(x) + O(x^7) .$$

(ii) $D(x) = -15x^4+900x^3+900x^2$ so the roots of $D(x)$ are

$x = 0$ (twice)

$x = -0.9839$

$x = 60.9839.$

The roots of $a_2(x)$ are $x = -0.8730, x = 6.8730$. Consideration of the proof of Theorem 1 in [2] shows that each of these roots corresponds to a pole on only one of the sheets

$$y(x) = \frac{-a_1(x)-x\sqrt{-15x^2+900x+900}}{2a_2(x)}$$

$$y_1(x) = \frac{-a_1(x)+x\sqrt{-15x^2+900x+900}}{2a_2(x)},$$

and a removable singularity on the other. A simple calculation or consideration of the later graphs shows that $y(x)$ has no poles and to ensure that $y(x)$ is single valued we simply need

to define its domain R as $\mathbf{C}\setminus\{x \in \mathbf{R} : x \in (-\infty, -0.9839) \text{ or } x \in (60.9839, \infty)\}$. Graphs of $y(z)$ and the error function $e(z) = y(z) - \log(1+z)$ are now presented. These graphs are over the region $\{x + \mathbf{i}y \in \mathbf{C} : |x| \leq 2, |y| \leq 2\}$ with mesh spacing of 0.1 using PC-Matlab. The point $-2 - 2\mathbf{i}$ corresponds to the lower left corner. It should be noted that some calculation error, particularly near $z = -1$ is inevitable but given that PC-Matlab uses double precision and that it is understood that we are working with an open mesh, this effect is minimal.

Fig.1 and Fig.2 are the real and imaginary parts respectively of the approximation. To the naked eye these surfaces are virtually indistinguishable from those of $\log(1+z)$ (although $\lim_{t\to -1} \log(1+t) = \infty \neq y(-1)$). Graphs of the real and imaginary parts of the error function (clearly not exact at $z = -1$) also follow. These are shown truncated at successively smaller values so as to illustrate the error behaviour throughout this region.

Fig	Real/Imag	Truncation interval
3	real(e(z))	$\pm\infty$
4	imag(e(z))	$\pm\infty$
5	real(e(z))	$\pm 10^{-1}$
6	imag(e(z))	$\pm 10^{-1}$
7	real(e(z))	$\pm 10^{-2}$
8	imag(e(z))	$\pm 10^{-2}$
9	real(e(z))	$\pm 10^{-3}$
10	imag(e(z))	$\pm 10^{-3}$
11	real(e(z))	$\pm 10^{-4}$
12	imag(e(z))	$\pm 10^{-4}$
13	real(e(z))	$\pm 10^{-5}$
14	imag(e(z))	$\pm 10^{-5}$

Fig.15 and Fig.16 are contour maps of the real and imaginary parts of $e(z)$ with contours drawn at $\{\pm 10^{-3}, \pm 10^{-4}, \pm 10^{-5}, \pm 10^{-6}, \pm 10^{-7}, \pm 10^{-8}\}$.

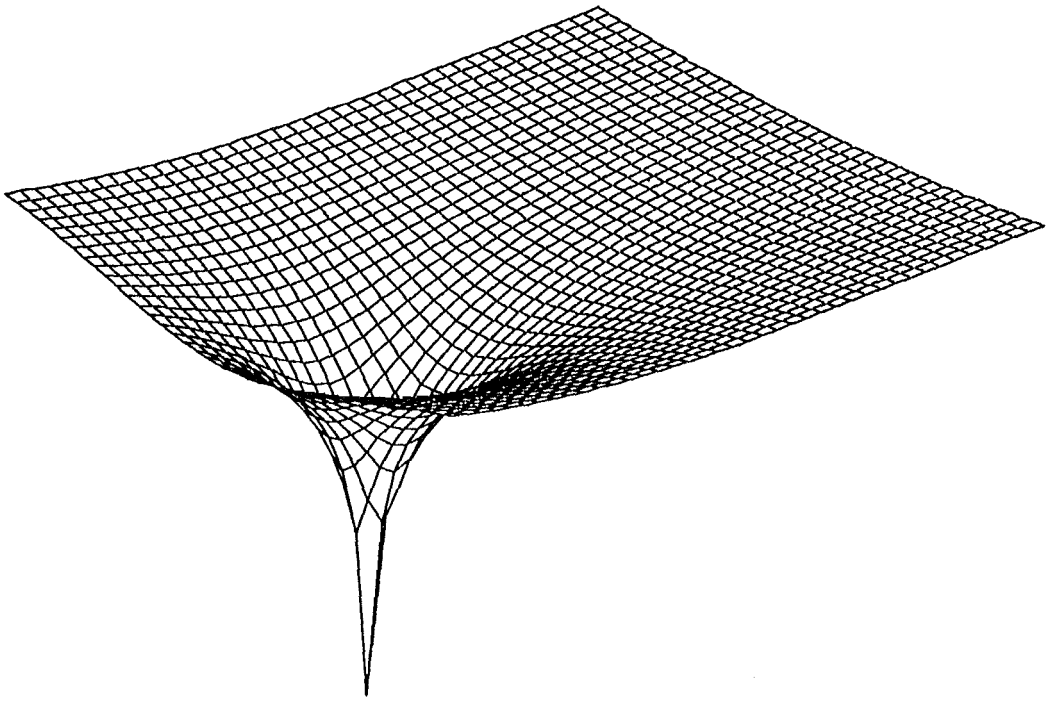


Figure 1 Real(y(z)).

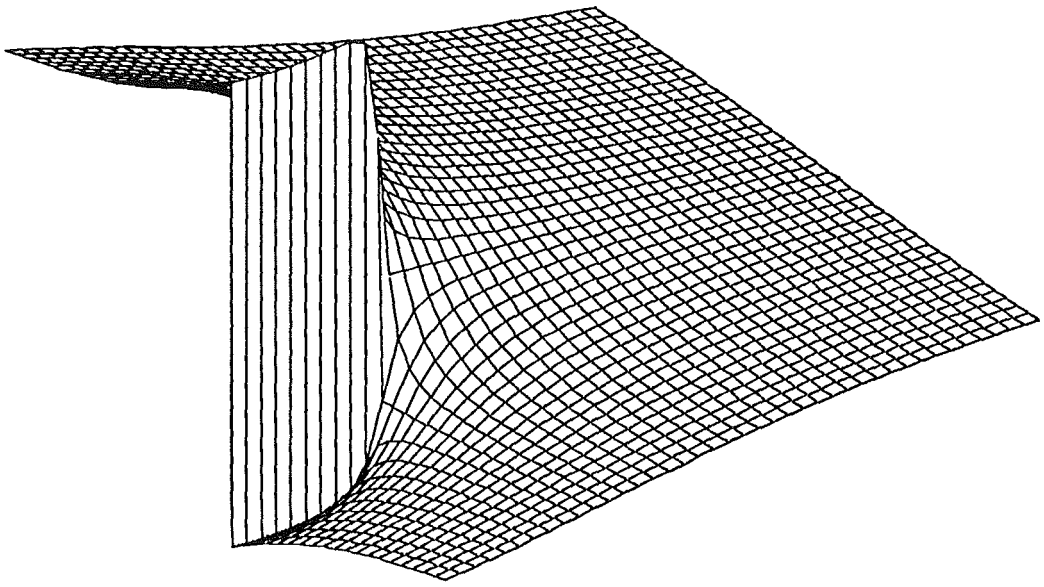


Figure 2 Imag(y(z)).

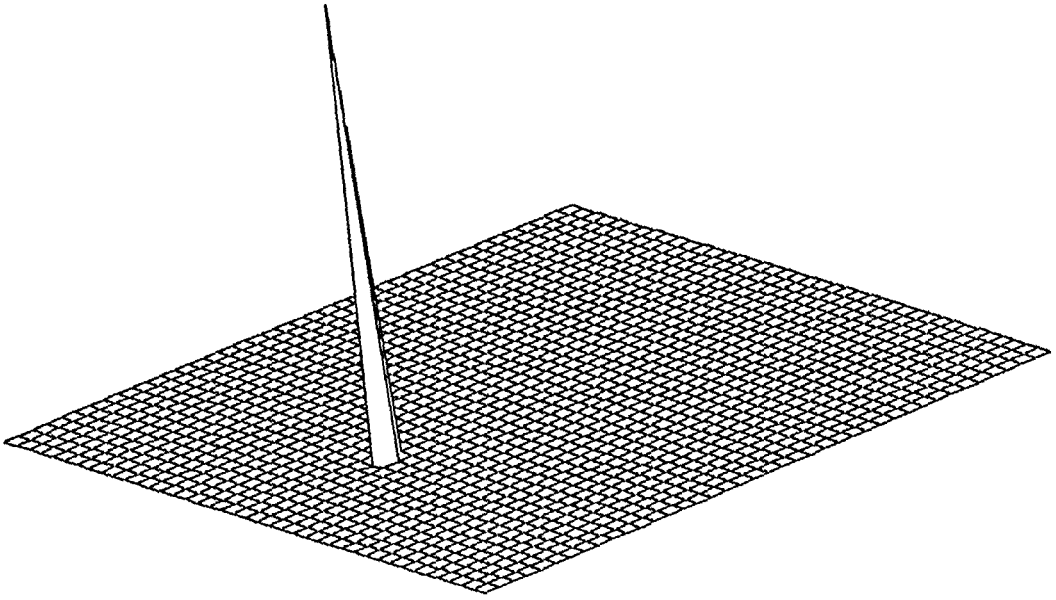


Figure 3 Real(e(z)). Truncation $\pm\infty$

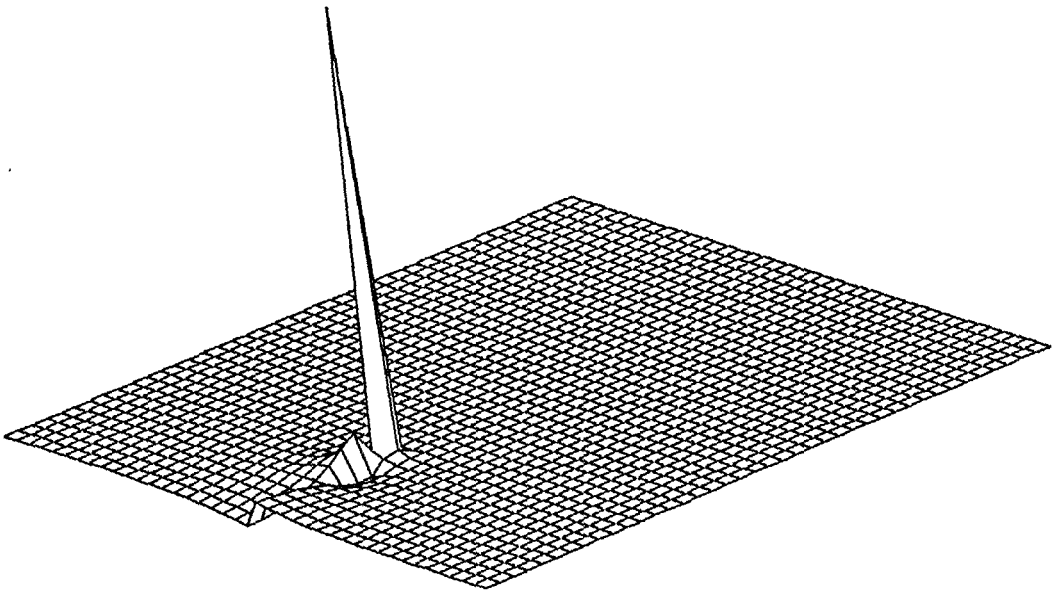


Figure 4 Imag(e(z)). Truncation $\pm\infty$

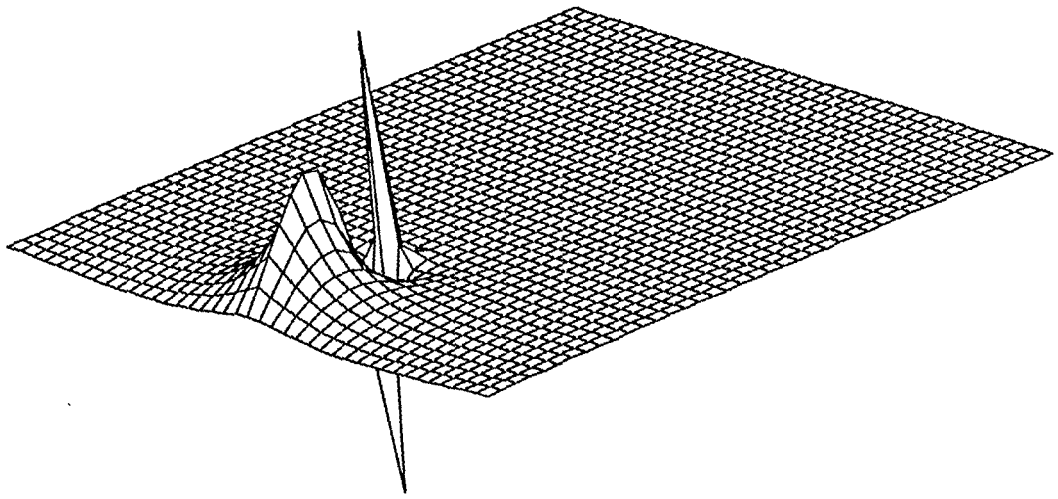


Figure 5 Real(e(z)). Truncation $\pm 10^{-1}$

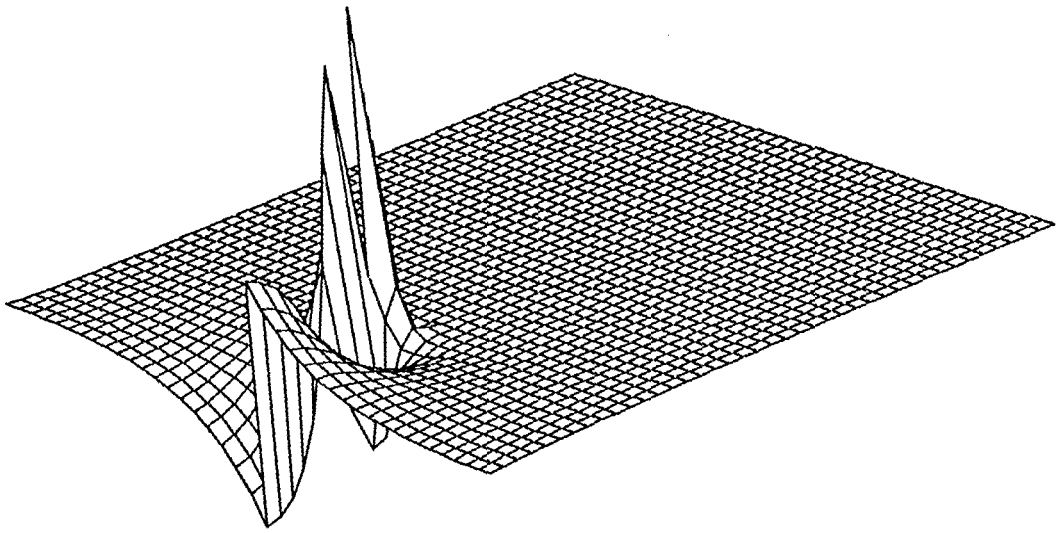


Figure 6 Imag(e(z)). Truncation $\pm 10^{-1}$

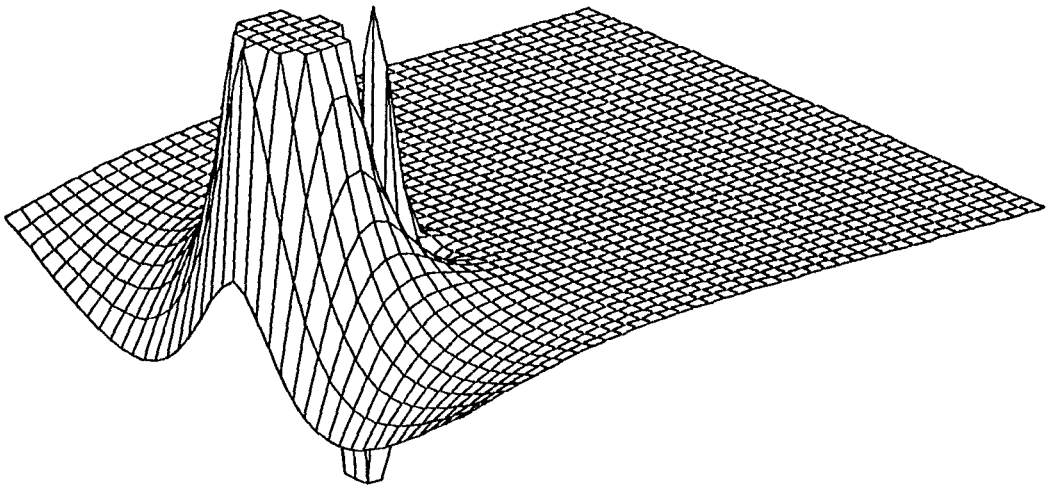


Figure 7 Real(e(z)). Truncation $\pm 10^{-2}$

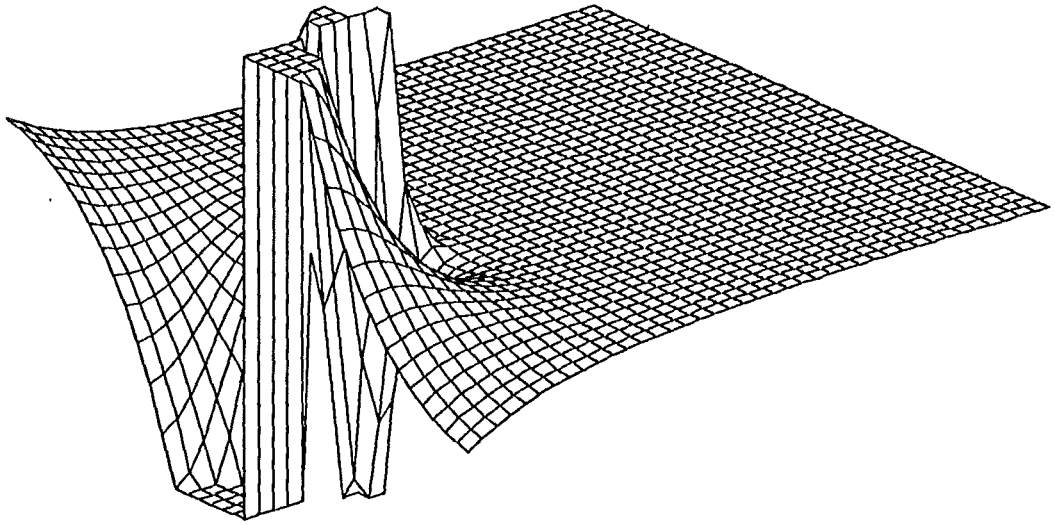


Figure 8 Imag(e(z)). Truncation $\pm 10^{-2}$

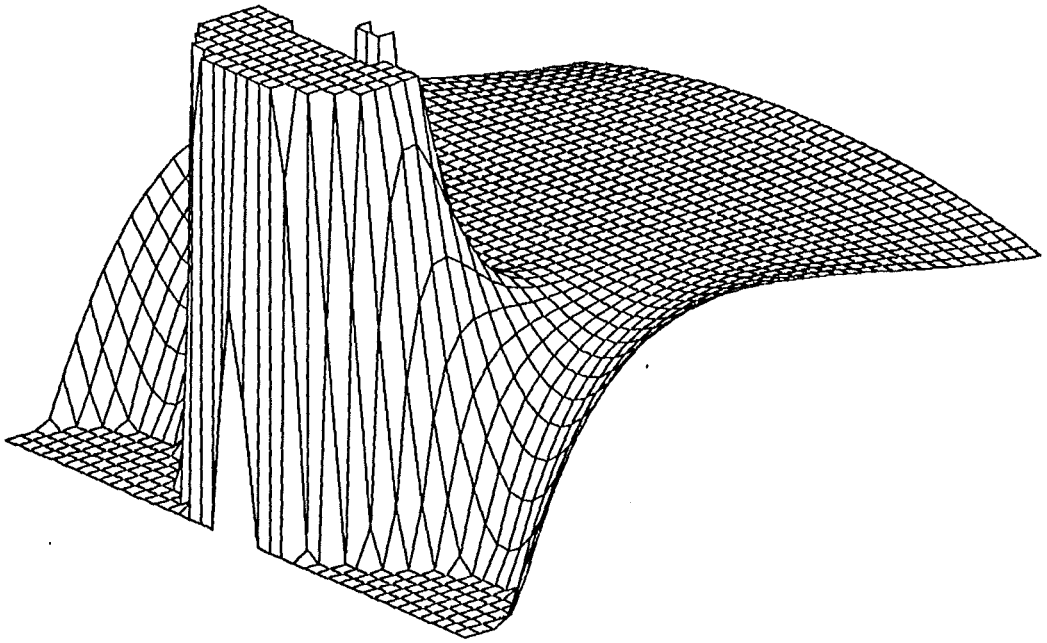


Figure 9 Real(e(z)). Truncation $\pm 10^{-3}$

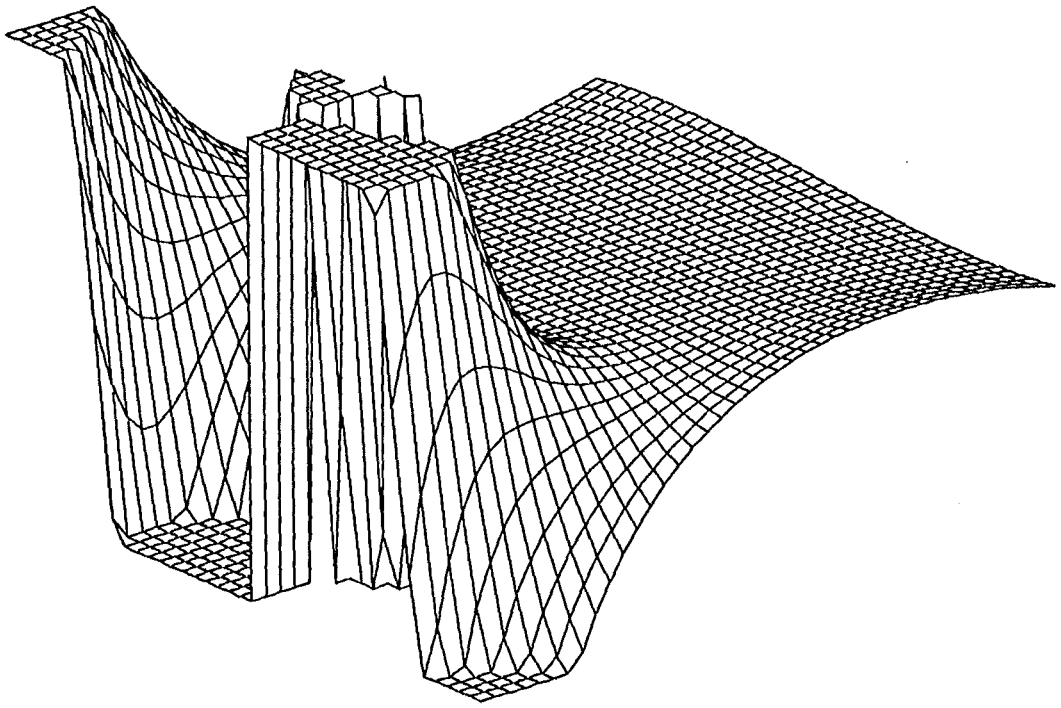


Figure 10 Imag(e(z)). Truncation $\pm 10^{-3}$

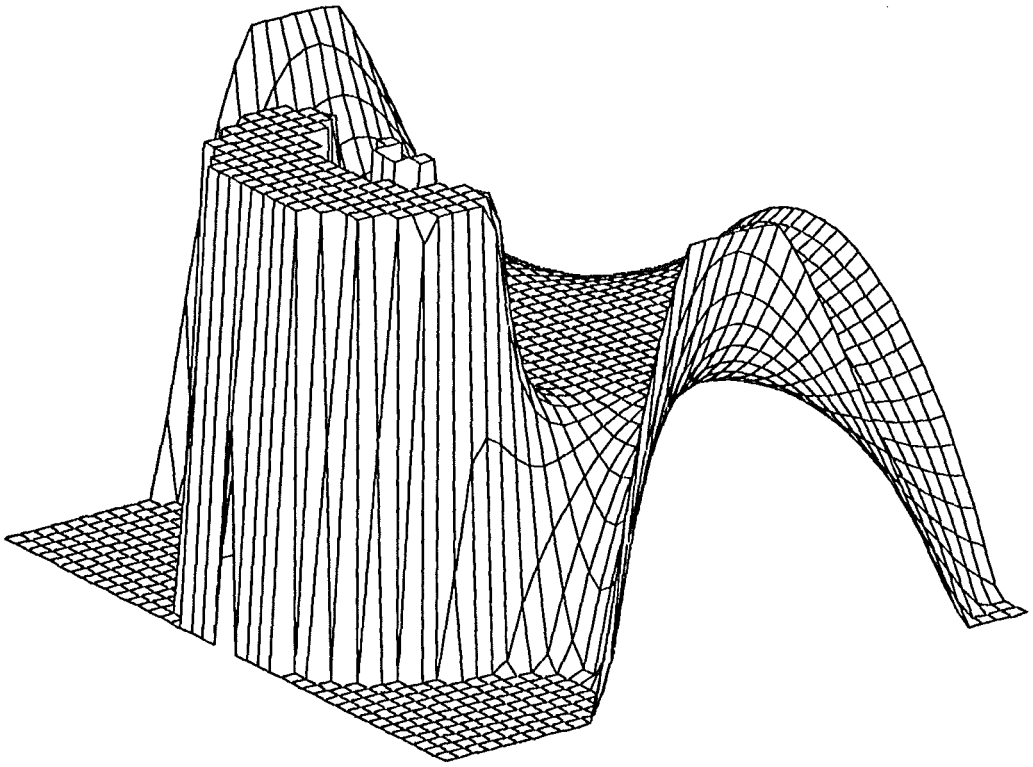


Figure 11 Real(e(z)). Truncation $\pm 10^{-4}$

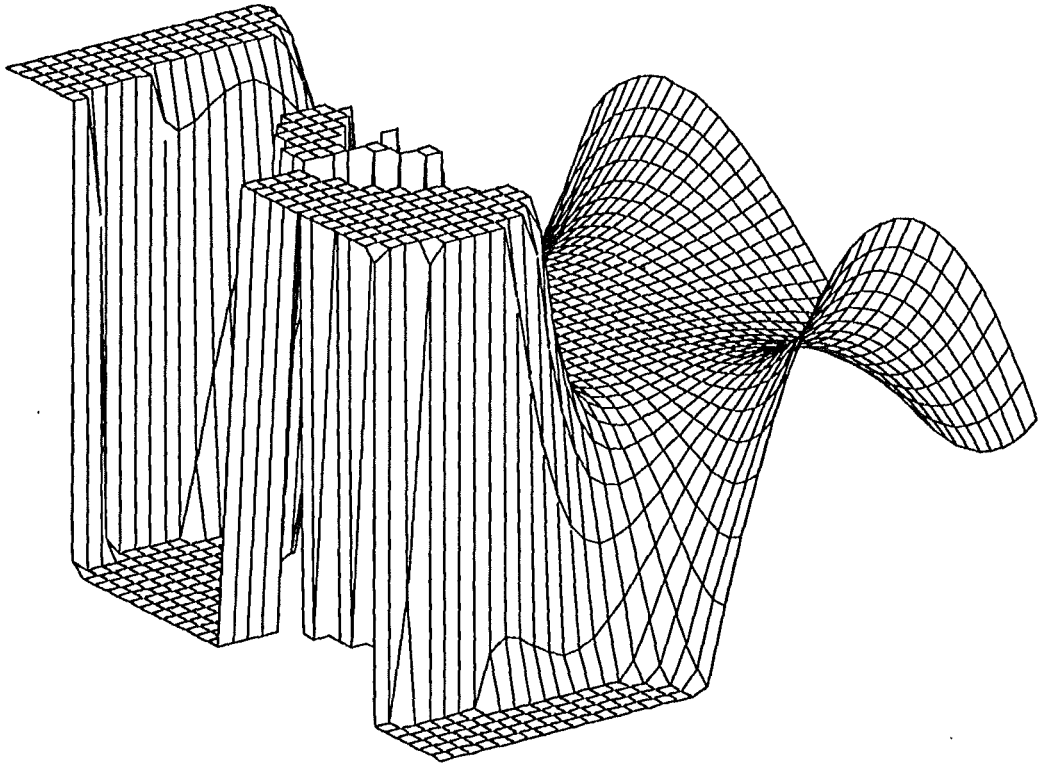


Figure 12 Imag(e(z)). Truncation $\pm 10^{-4}$

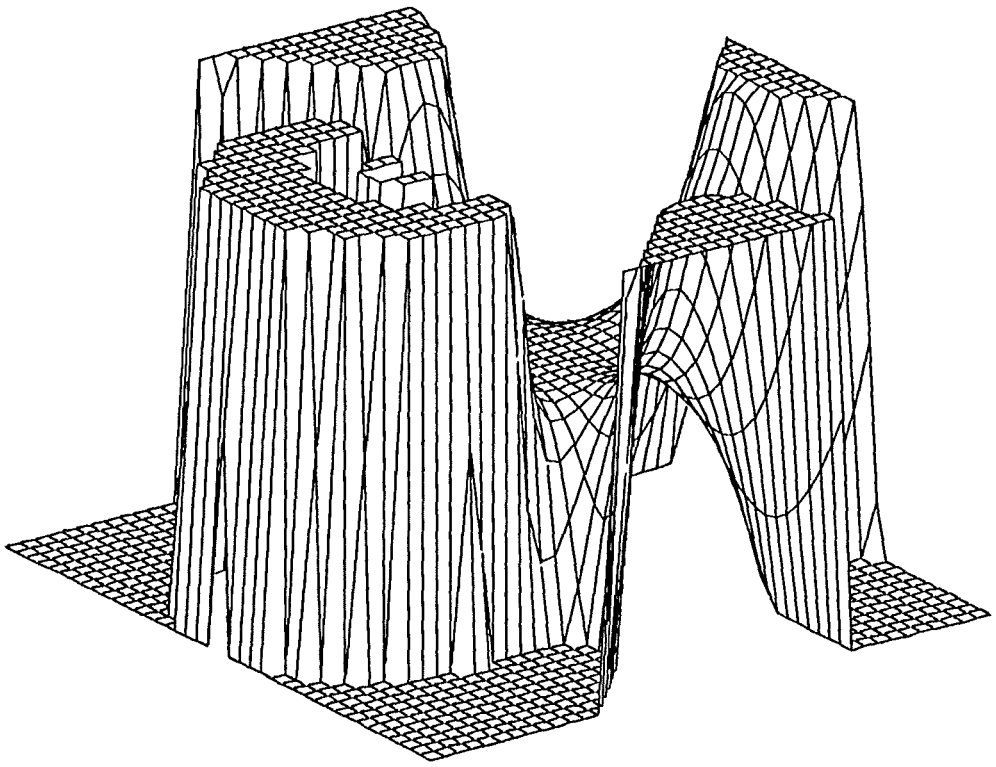


Figure 13 Real(e(z)). Truncation $\pm 10^{-5}$

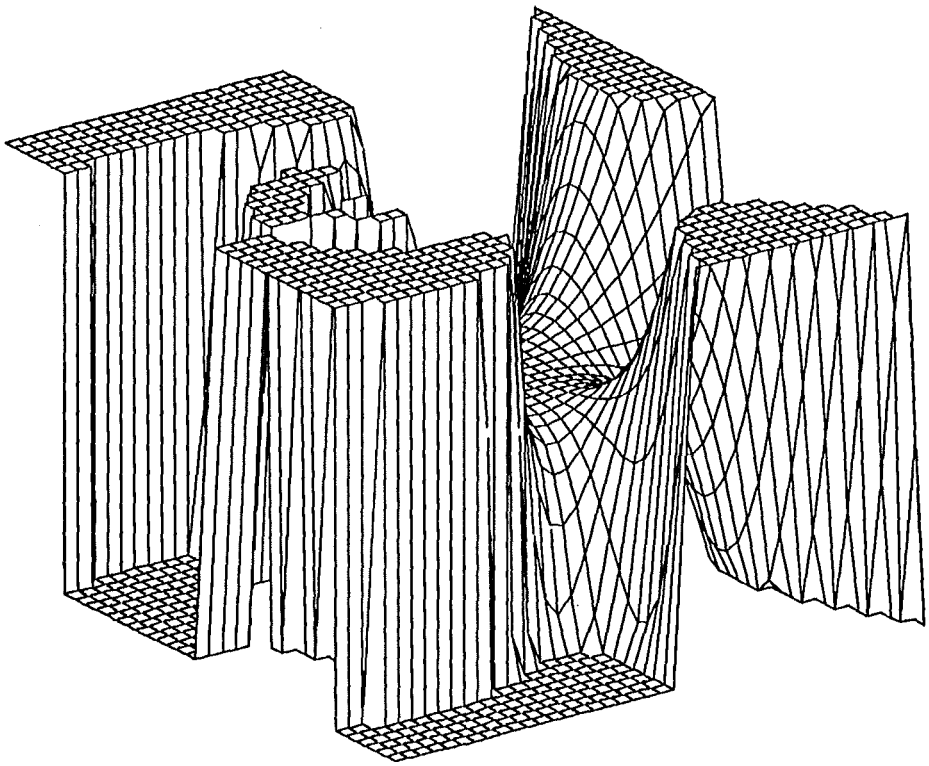


Figure 14 Imag(e(z)). Truncation $\pm 10^{-5}$

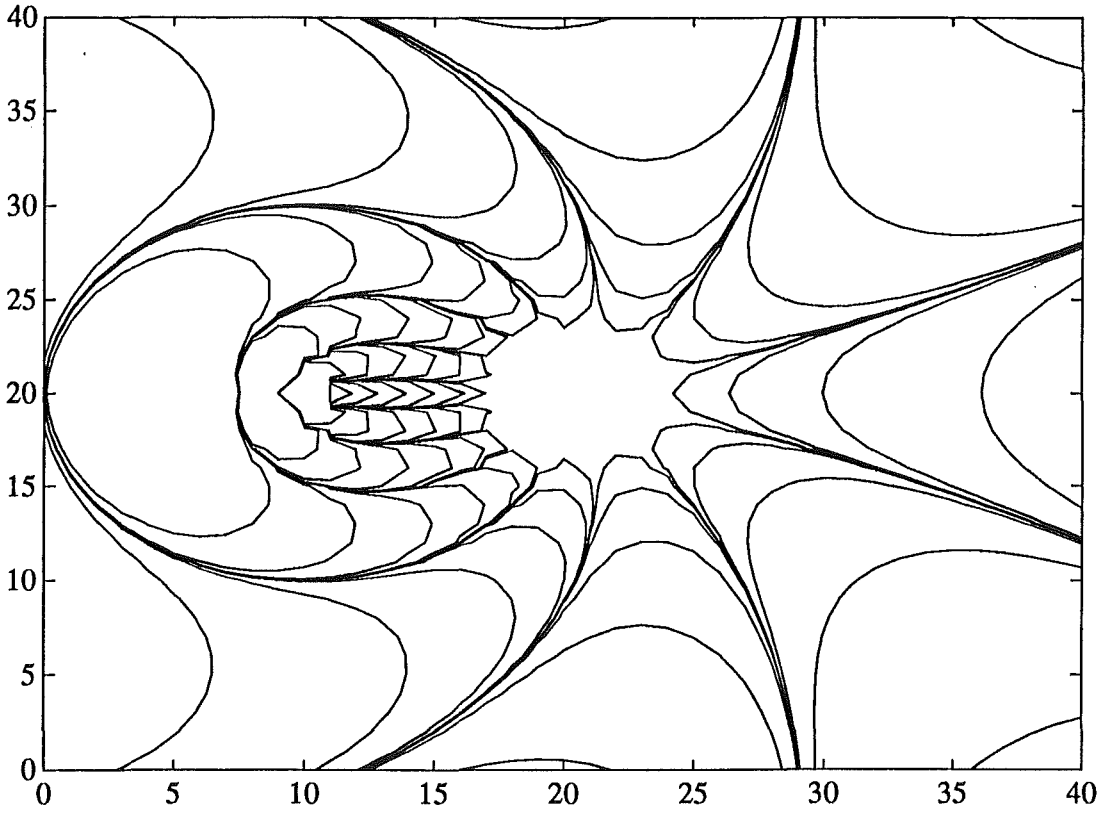


Figure 15 Contour map of Real(e(z)).

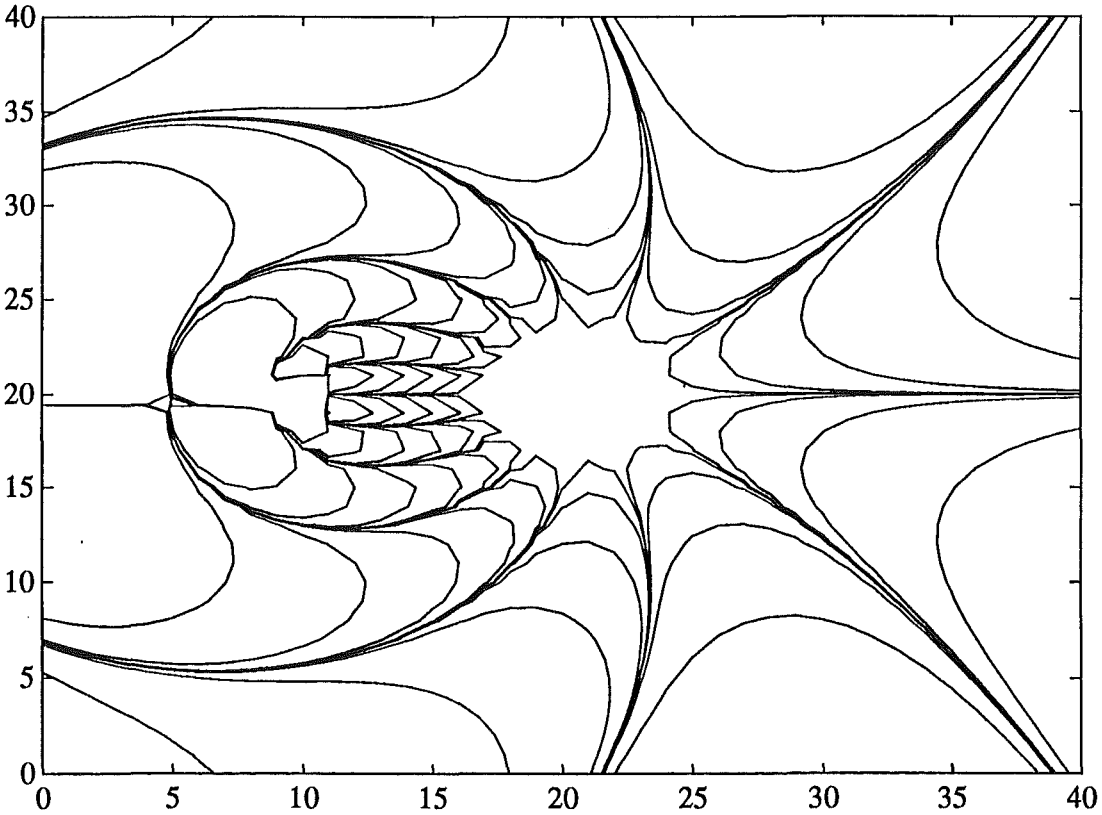


Figure 16 Contour map of Imag(e(z)).

Comparison to a Padé approximation and to a Taylor polynomial.

In order to gauge whether the quadratic approximation is worthwhile it is useful to compare its performance with those of the Padé and Taylor approximations which match $f(x)$ to the same order. Here we choose $p(x)$, the $(3,3)$ Padé approximation and $t(x)$, the Taylor polynomial of degree 6.

Note that

$$\begin{aligned} y(x) &= f(x) + O(x^7) \\ p(x) &= f(x) + O(x^7) \\ t(x) &= f(x) + O(x^7) \ . \end{aligned}$$

The (3,3) Padé approximation to log(1+x). The $(3,3)$ Padé approximation to $\log(1+x)$ is given by

$$p(x) = \frac{11x^3 + 60x^2 + 60x}{3x^3 + 36x^2 + 90x + 60}$$

$p(x)$ has poles at $x = -8.87, -2.00, -1.13$.

Graphs of the real and imaginary parts of $p(z)$ follow. Because of the obvious difficulties, $p(-2)$ has been set to zero.

Fig.17 and Fig.18 are the real and imaginary parts respectively of $p(z)$.

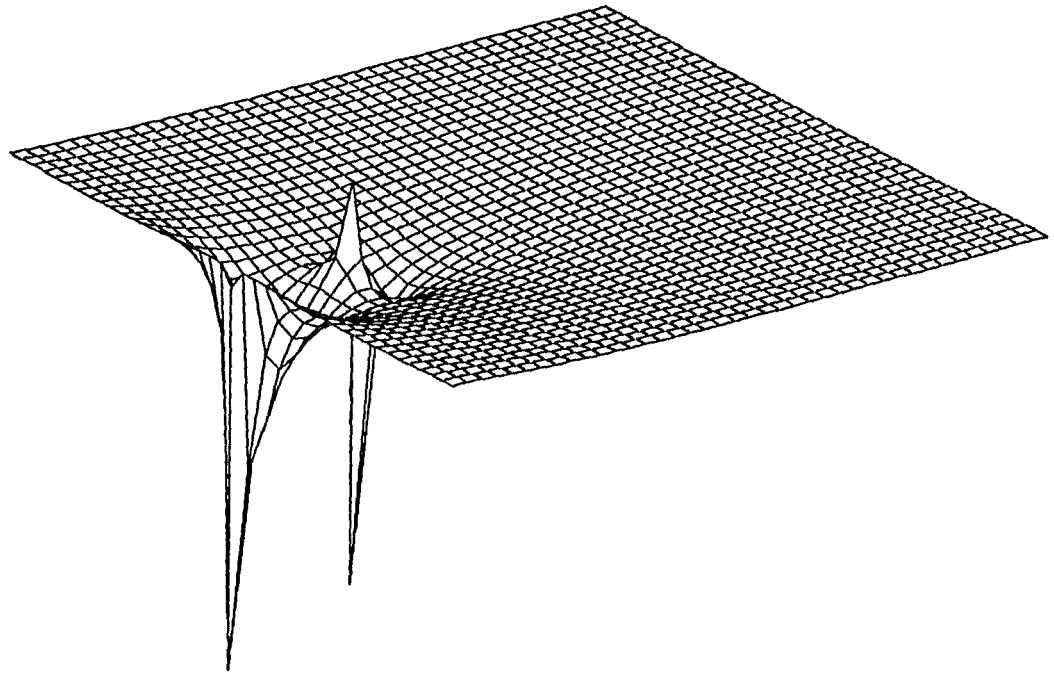


Figure 17 Real(p(z)).

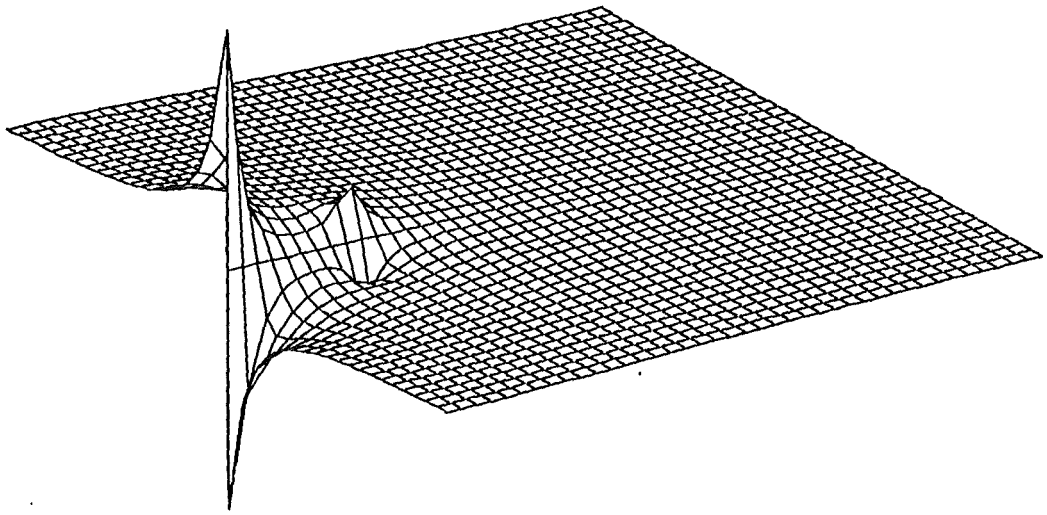


Figure 18 Imag(p(z)).

Graphs of the real and imaginary parts of the error function $e(z) = p(z) - \log(1+z)$ follow.

Fig	Real/Imag	Truncation interval
19	real(e(z))	$\pm\infty$
20	imag(e(z))	$\pm\infty$
21	real(e(z))	$\pm 10^{-1}$
22	imag(e(z))	$\pm 10^{-1}$
23	real(e(z))	$\pm 10^{-2}$
24	imag(e(z))	$\pm 10^{-2}$
25	real(e(z))	$\pm 10^{-3}$
26	imag(e(z))	$\pm 10^{-3}$

Figures 27 and 28 are contour maps of Real$(e(z))$ and Imag$(e(z))$ with contours drawn at $\{\pm 10^{-3}, \pm 10^{-4}, \pm 10^{-5}, \pm 10^{-6}, \pm 10^{-7}, \pm 10^{-8}\}$ as in the previous case.

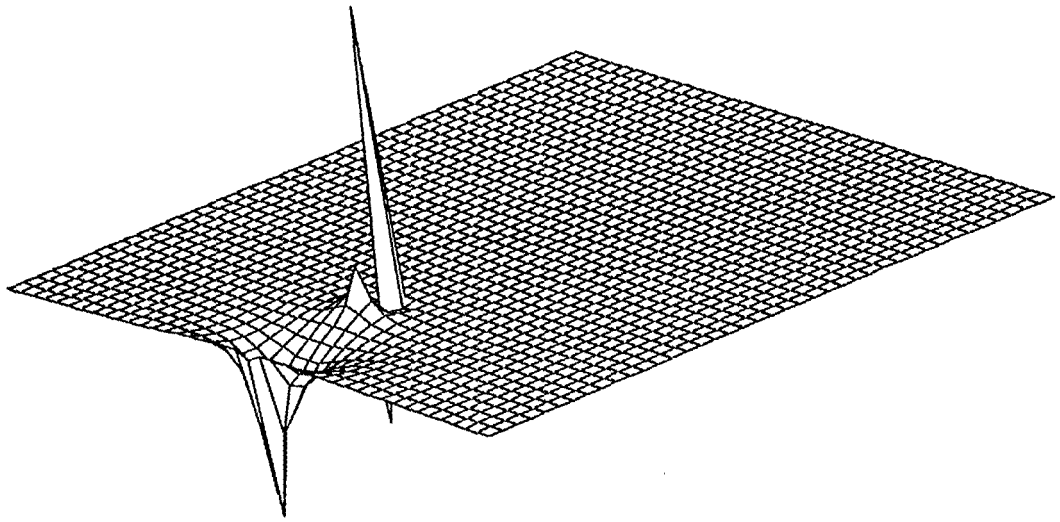


Figure 19 Real(e(z)). Truncation $\pm\infty$

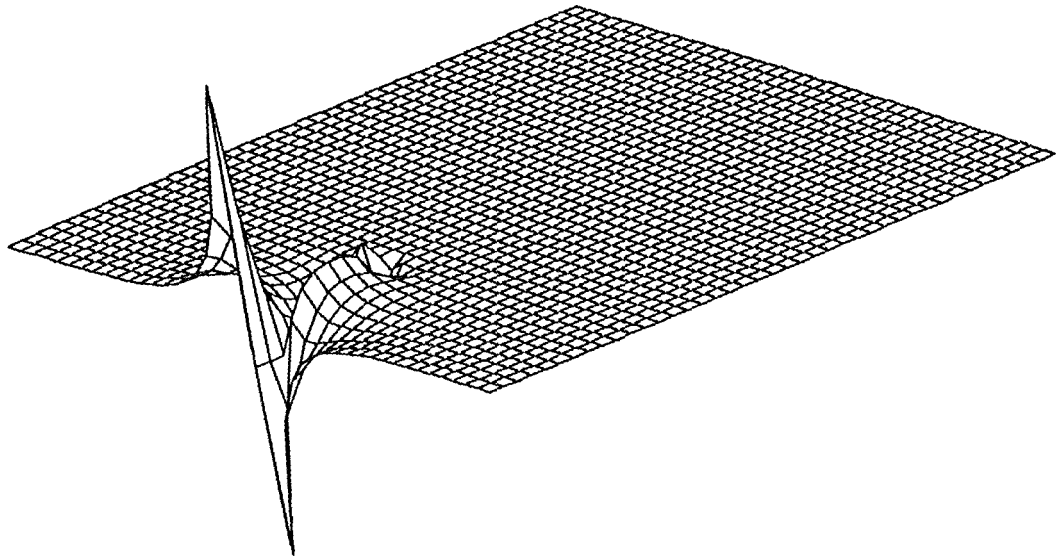


Figure 20 Imag(e(z)). Truncation $\pm\infty$

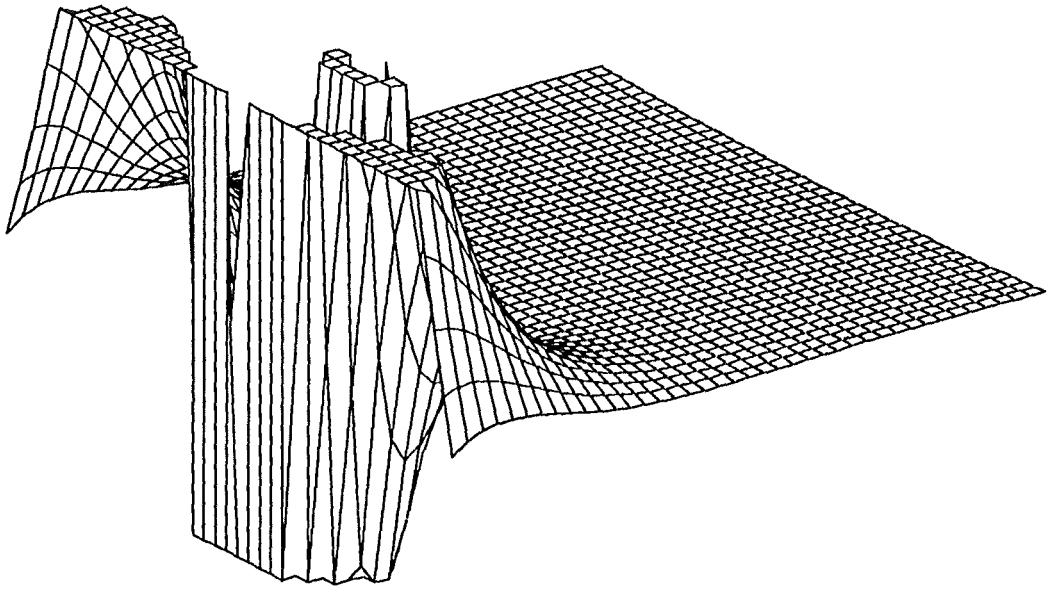


Figure 21 Real(e(z)). Truncation $\pm 10^{-1}$

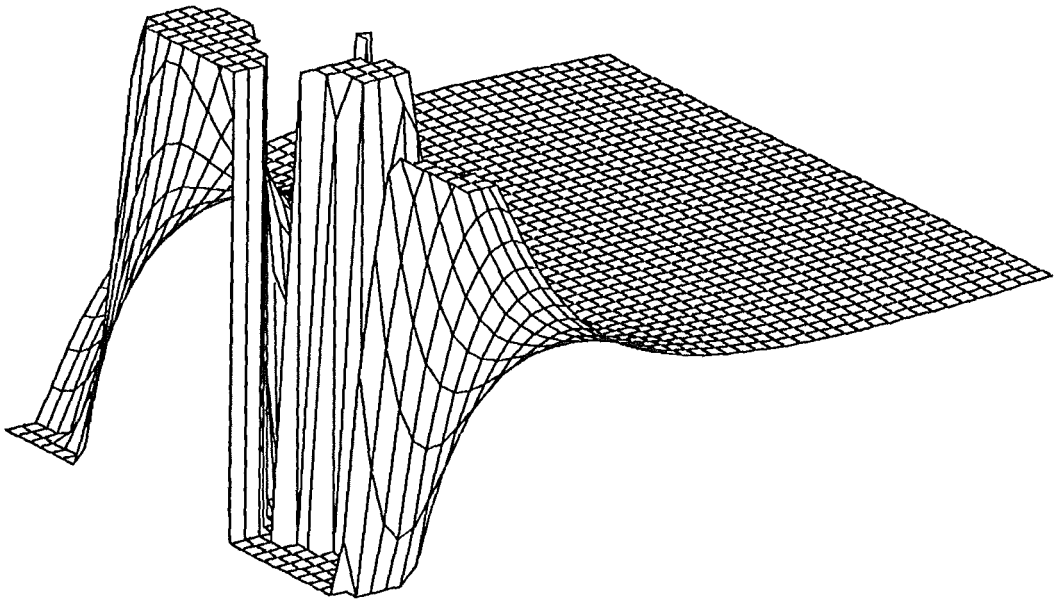


Figure 22 Imag(e(z)). Truncation $\pm 10^{-1}$

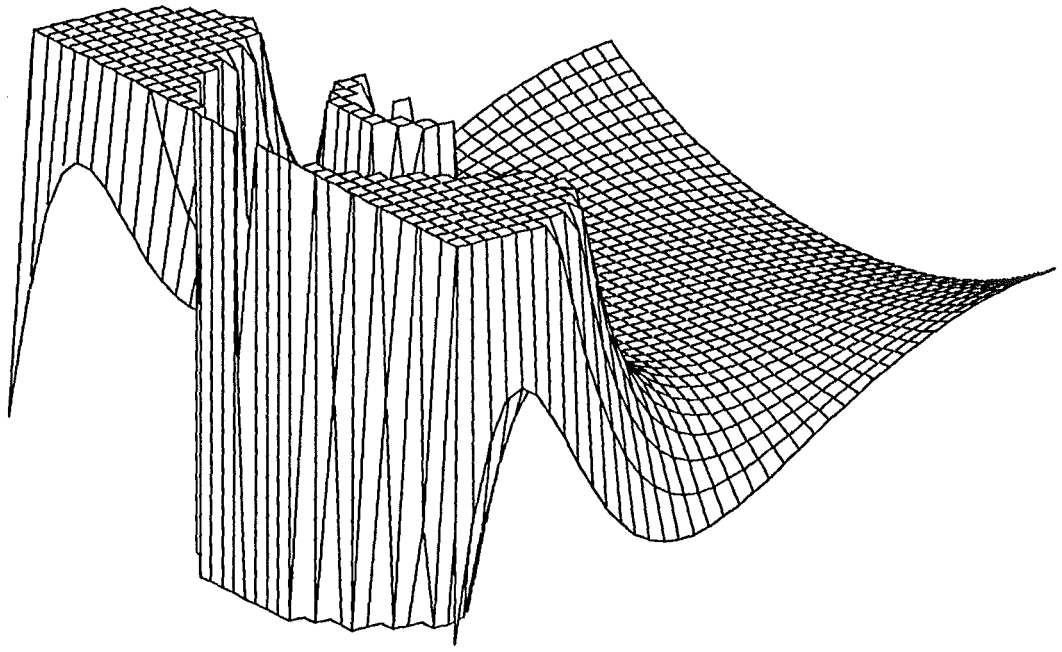


Figure 23 Real(e(z)). Truncation $\pm 10^{-2}$

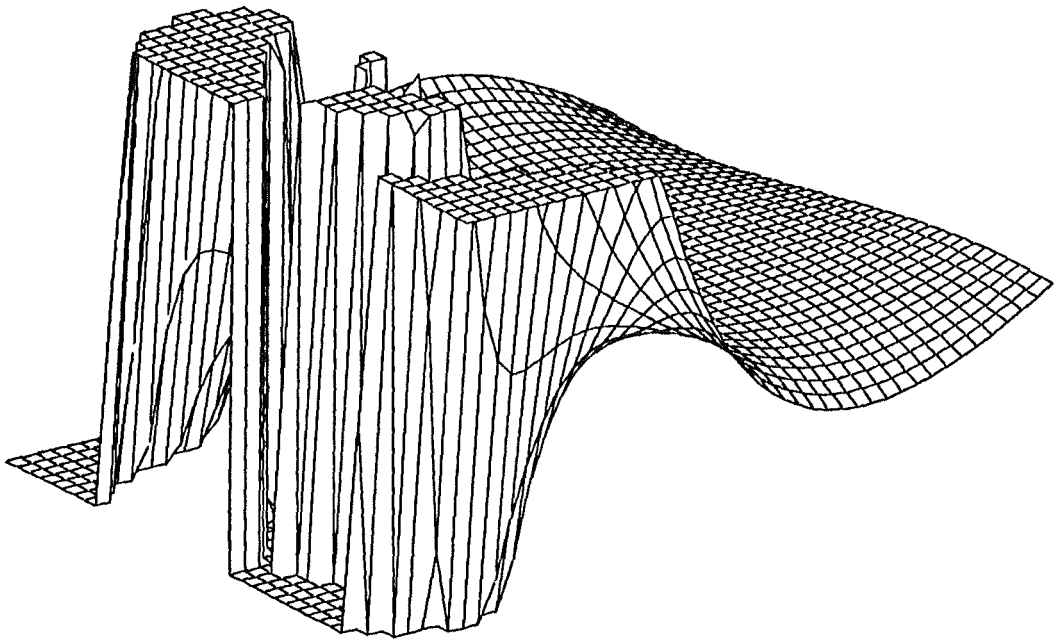


Figure 24 Imag(e(z)). Truncation $\pm 10^{-2}$

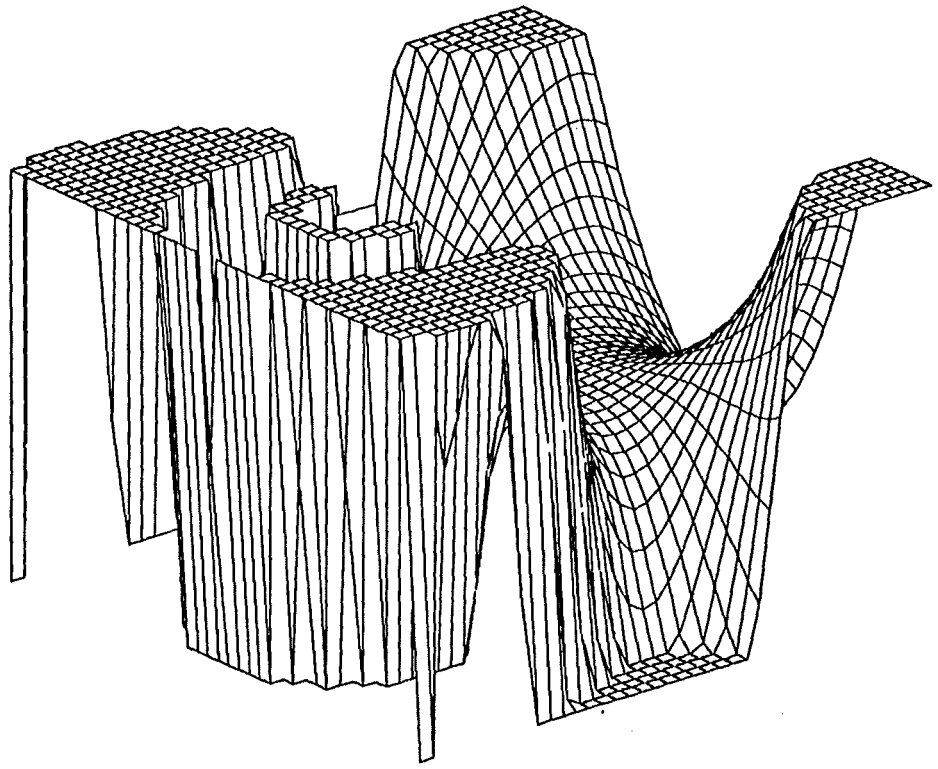


Figure 25 Real(e(z)). Truncation $\pm 10^{-3}$

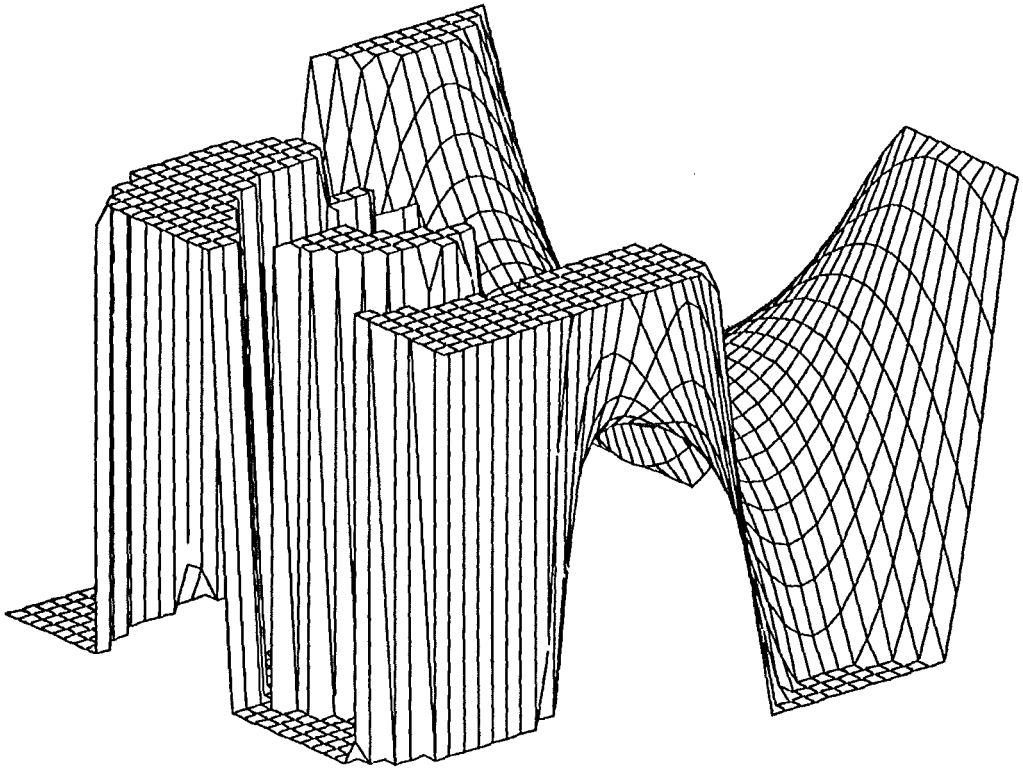


Figure 26 Imag(e(z)). Truncation $\pm 10^{-3}$

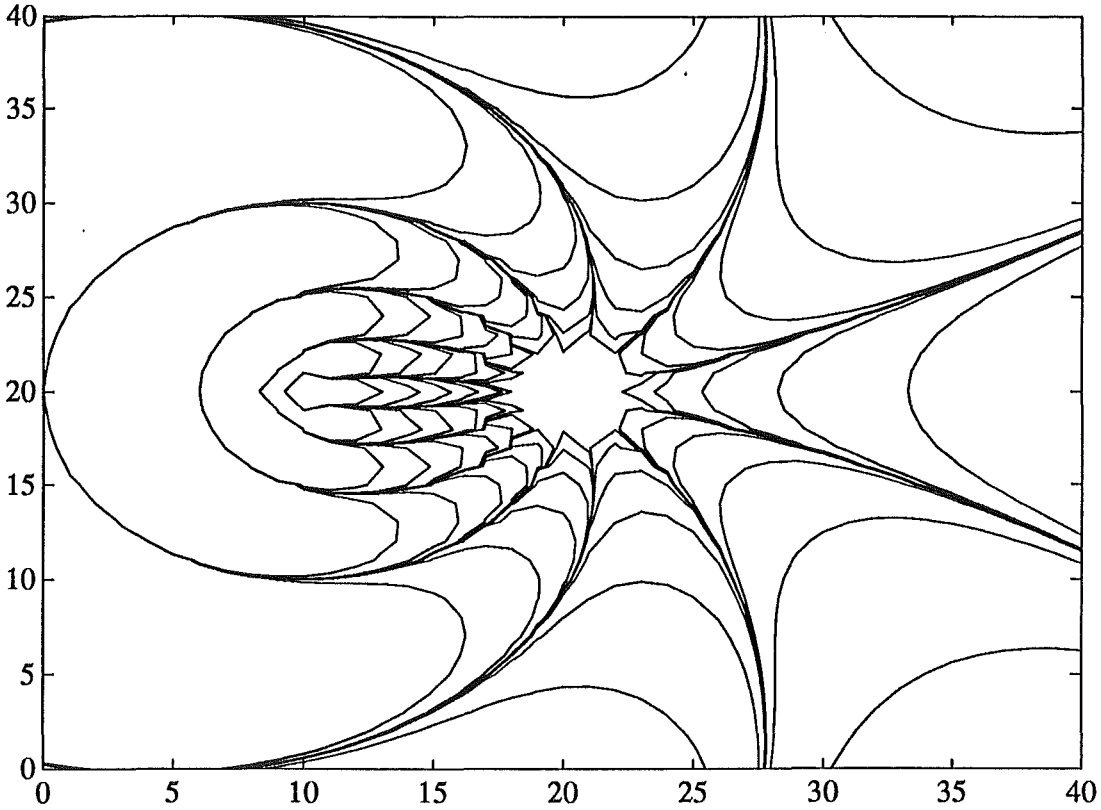


Figure 27 Contour map of Real(e(z)).

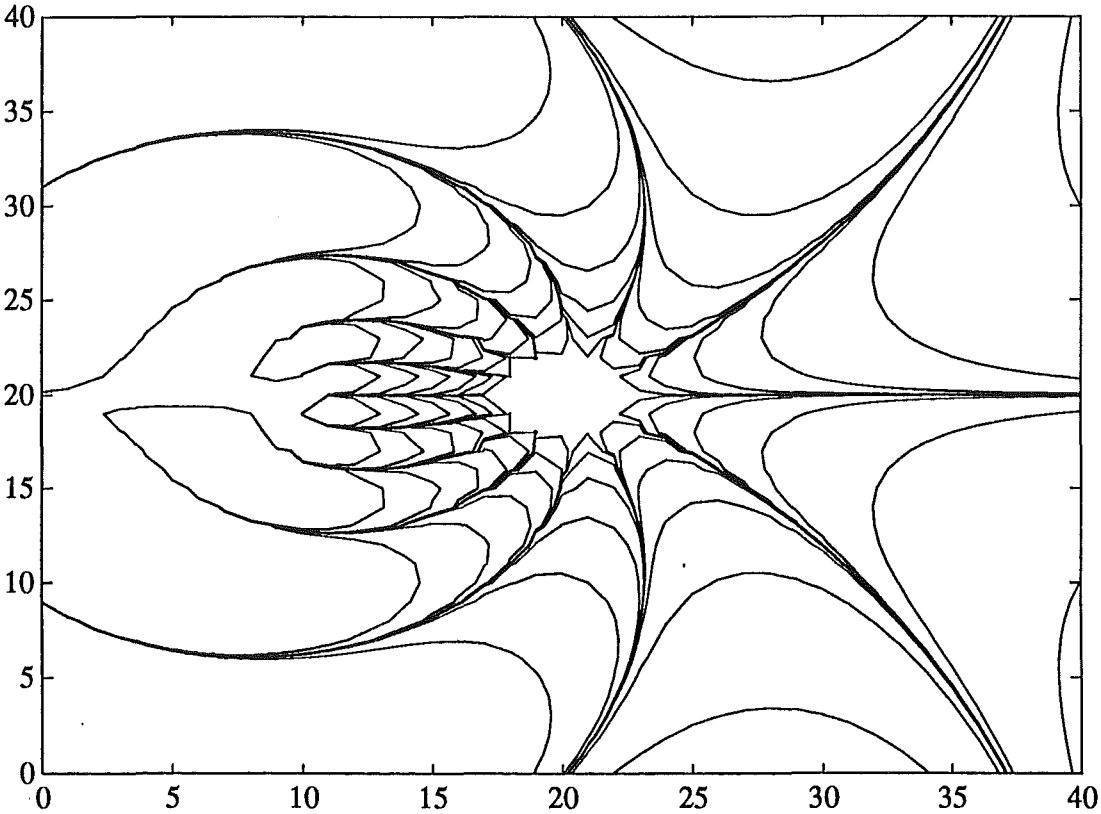


Figure 28 Contour map of Imag(e(z)).

Consideration of these pictures makes it clear that $y(z)$ (the quadratic function) is a considerably better approximation to $f(z)$ than $p(z)$ (the rational function) both around the origin and the branch point. Clearly the branch point structure of $y(z)$ is very similar to that of $f(z)$.

The degree 6 Taylor polynomial approximation to log(1+x). Finally graphs of $t(z)$, the Taylor polynomial of degree 6 and the error function $e(z) = t(z) - \log(1+z)$ are given. $t(z)$ is clearly inferior to both $y(z)$ and $p(z)$ as an approximation.

Figures 29 and 30 are the real and imaginary parts of $t(z)$.

Figures 31 and 32 are the real and imaginary parts of $e(z)$ truncated at $\pm 10^{-1}$.

Figures 33 and 34 are contour maps of Real$(e(z))$, Imag$(e(z))$ with contours drawn at $\{\pm 10^{-3}, \pm 10^{-4}, \pm 10^{-5}, \pm 10^{-6}, \pm 10^{-7}, \pm 10^{-8}\}$ as in the previous cases.

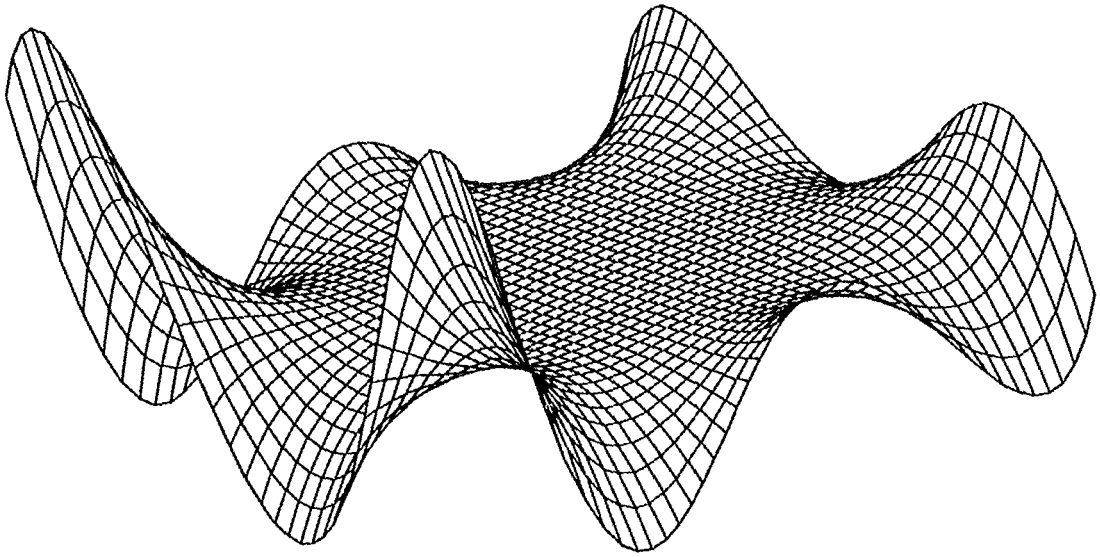


Figure 29 Real(t(z)).

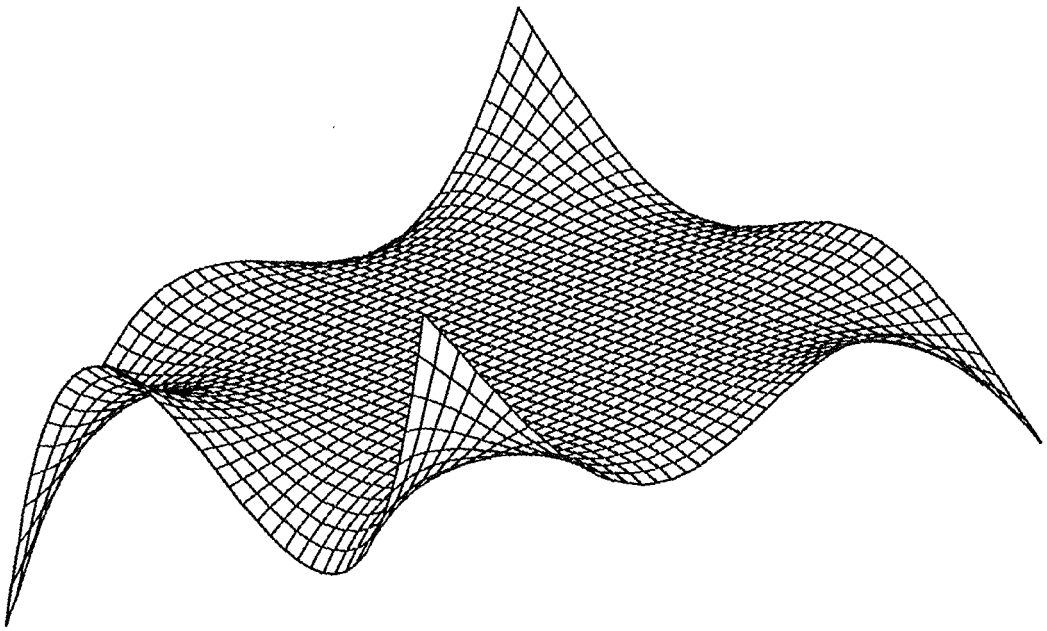


Figure 30 Imag(t(z)).

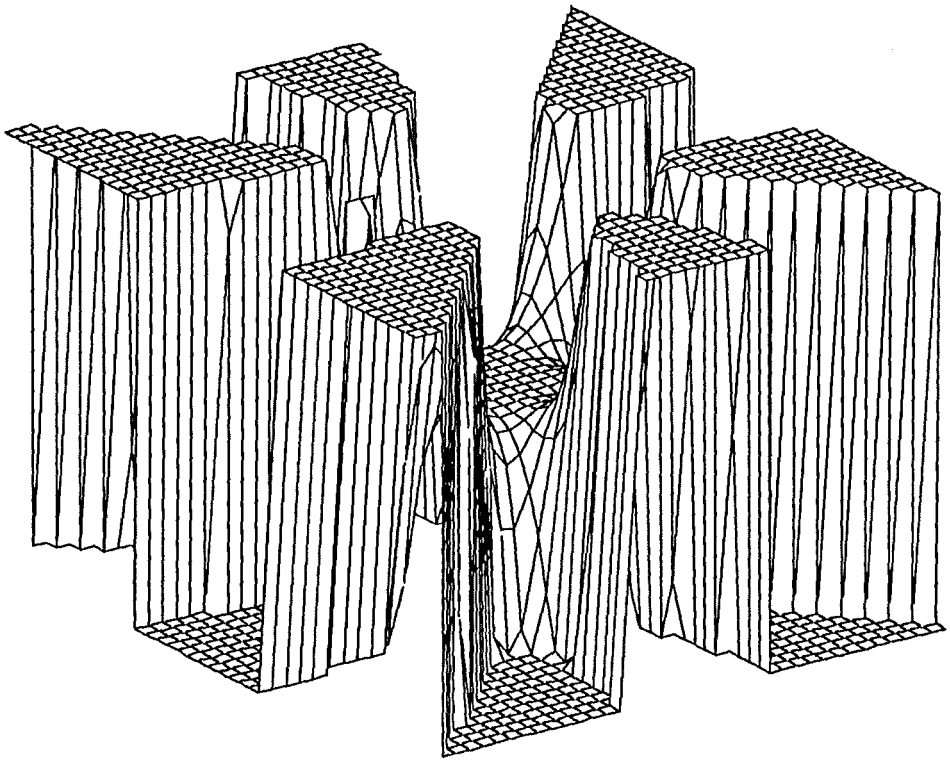


Figure 31 Real(e(z)). Truncation $\pm 10^{-1}$

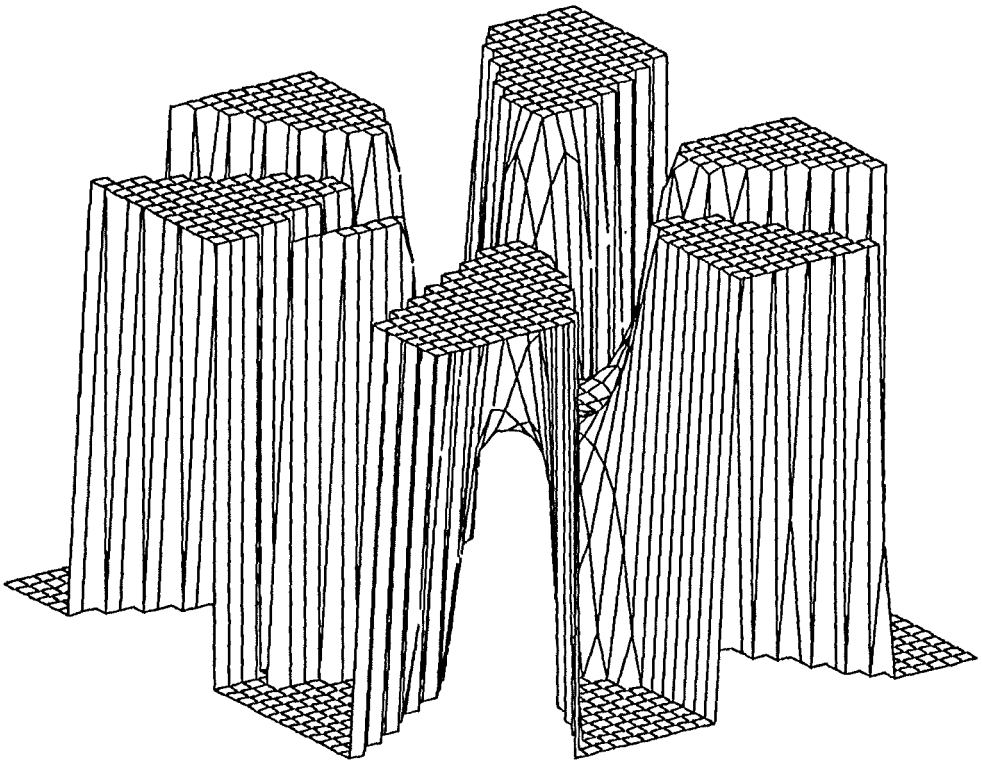


Figure 32 Imag(e(z)). Truncation $\pm 10^{-1}$

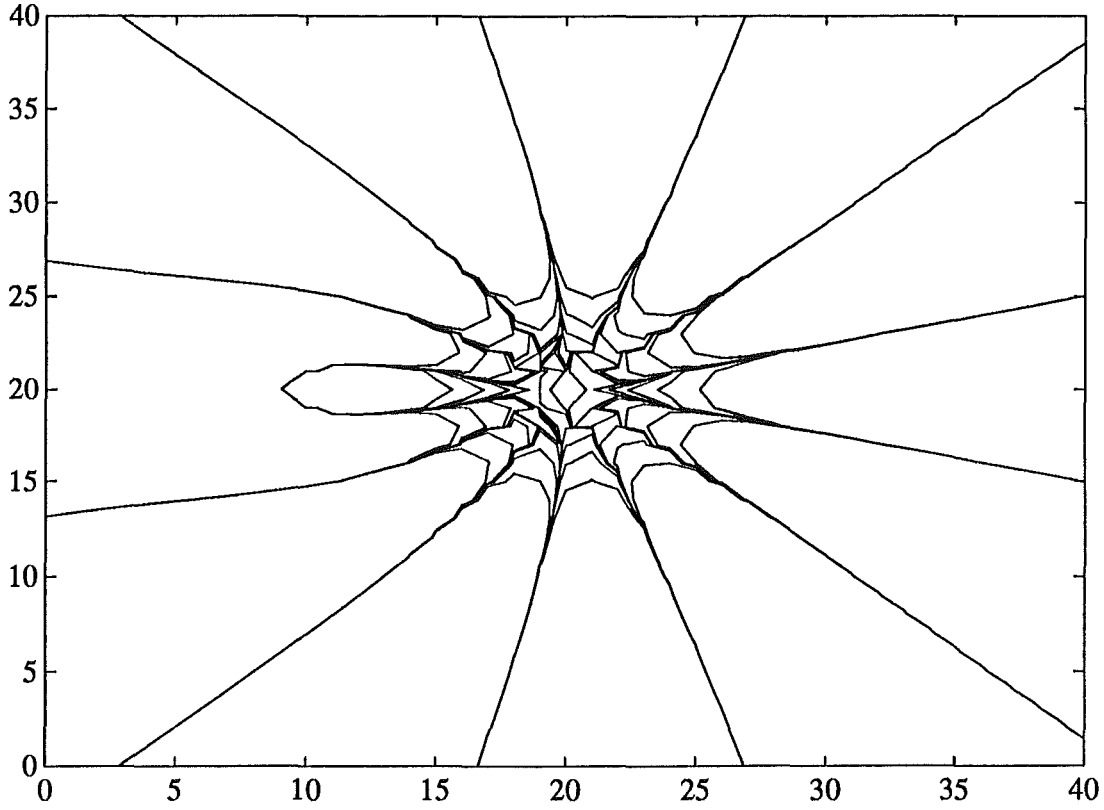


Figure 33 Contour map of Real(e(z)).

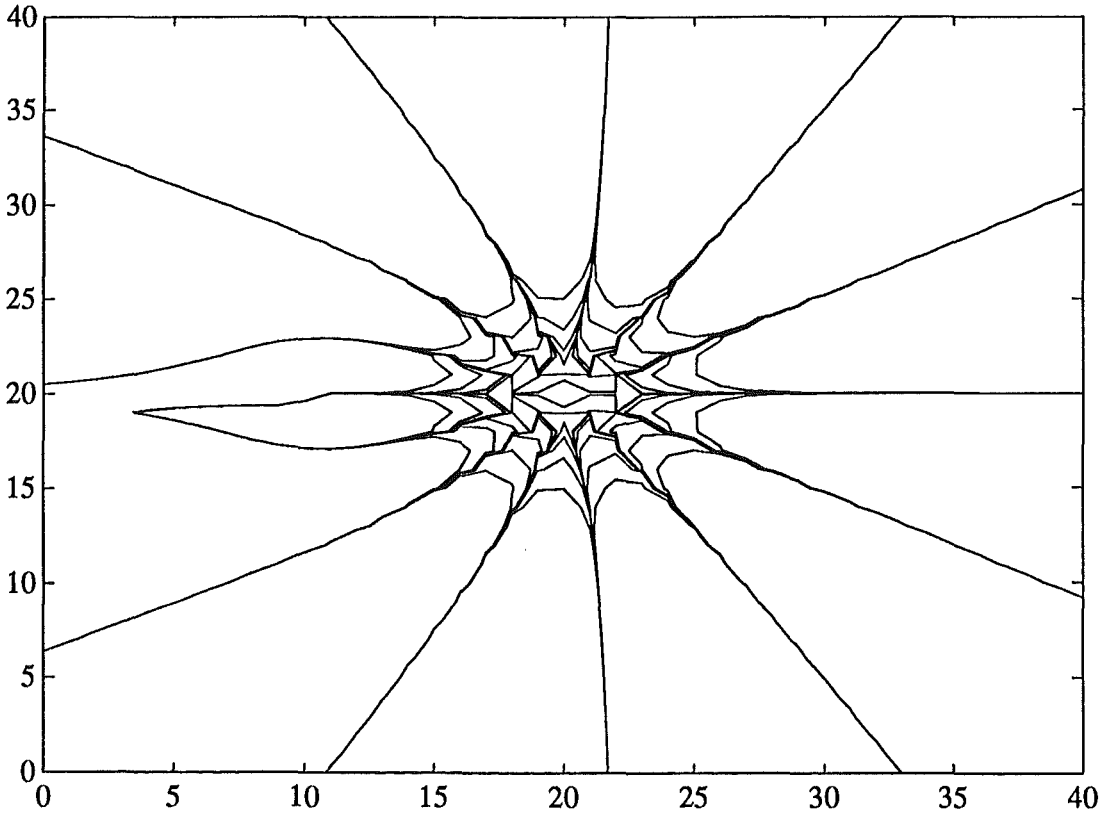


Figure 34 Contour map of Imag(e(z)).

3.2 Example 2

The (4,4,4) approximation to log(1+x). Note that :

(i)

$$\begin{aligned}&\left(6x^4-360x^3+180x^2+1080x+540\right)f(x)^2\\&+\left(-75x^4+1620x^3+5310x^2+3540x\right)f(x)\\&+260x^4-4080x^3-4080x^2\qquad\qquad=O\left(x^{14}\right)\end{aligned}$$

(ii)

$$\begin{aligned}D(x)=&-615x^8+229320x^7-4136580x^6\\&+12612600x^5+59667300x^4\\&+64033200x^3+21344400x^2\,.\end{aligned}$$

Let $d(x)=D(x)/x^2$. Using the results of Chapter 3 the approximation is :

$$y(x)=\frac{-a_1(x)+x\sqrt{d(x)}}{2a_2(x)}\,.$$

The roots of $d(x)$ are :

$$\begin{aligned}x&=354.0459\\x&=10.8301\pm0.06444\mathbf{i}\\x&=-0.9155\pm0.0005\mathbf{i}\\x&=-0.9972\end{aligned}$$

while the roots of $a_2(x)$ are :

$$\begin{aligned}x&=59.4440\\x&=2.2299\\x&=-0.6904\\x&=-0.9835\,.\end{aligned}$$

To ensure that $y(x)$ is single valued, cuts must be taken from the roots of $d(x)$. There are, of course, an infinite number of ways in which this may be done. It turns out that the simplest method gives a very good approximation to $\log(1+x)$. In the region presently under consideration $d(x)$ has 3 zeros. $x=-0.9972$ is close to the known branch point of $f(x)$ at $x=-1$ and this point is treated as in the previous example by taking $(\infty,-0.9972]$ as a cut. The other zeros are the conjugate pair $x=-0.9155\pm0.0005\mathbf{i}$. A cut could be taken from each point towards ∞ but in view of the fact that $f(x)$ is analytic on $\mathbf{C}\backslash\{x\in\mathbf{R}\,:\,x\in(-\infty,-1]\}$ this choice cannot be expected to give a good approximation. The other alternative is to take

a cut between the two points. In fact the behaviour of the two possible values of $y(x)$ on the real axis close to these points (Fig.35) suggests that the best choice is to take the cut $\{x + \mathbf{i}y \in \mathbf{C} : x = -0.9155,\ |y| \leq 0.0005\}$. It is of interest to note that this choice of cuts ensures that none of the roots of $a_2(x)$ are poles of $y(x)$.

Fig.35 shows the real parts of both possible continuations of $y(x)$ along the real axis; one, $y_+(x)$, denoted by a solid line, the other $y_-(x)$, by "*". $y_+(x)$, which is the analytic continuation of $y(x)$ along the real axis, has a pole at $z = -0.9835$. The effect of taking the cut between the points $z = -0.9155 \pm 0.0005\mathbf{i}$ is to "jump" from $y_+(x)$ to $y_-(x)$ at $x = -0.9155$. Fig.36 shows real(log(1 + x)) (although real(log(0)) $= -\infty$) and Fig.37 shows simultaneously real(log(1 + x)) and real($y(x)$) our approximation.

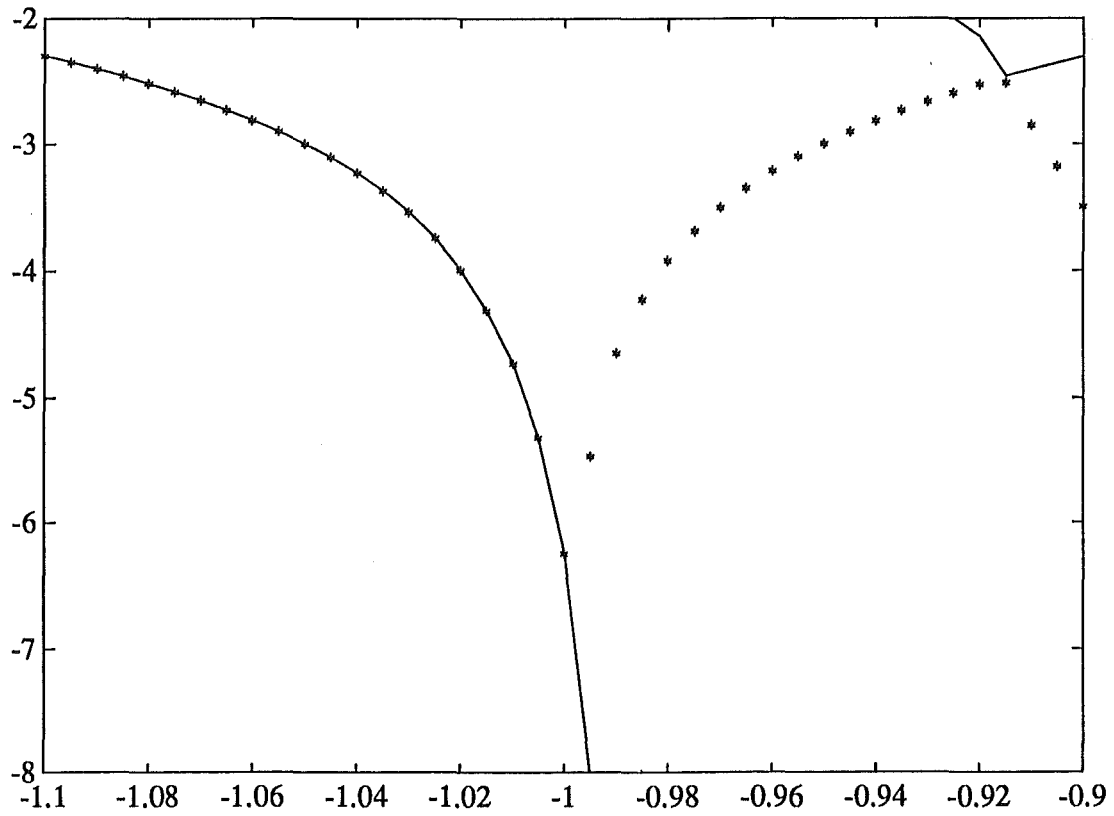


Figure 35

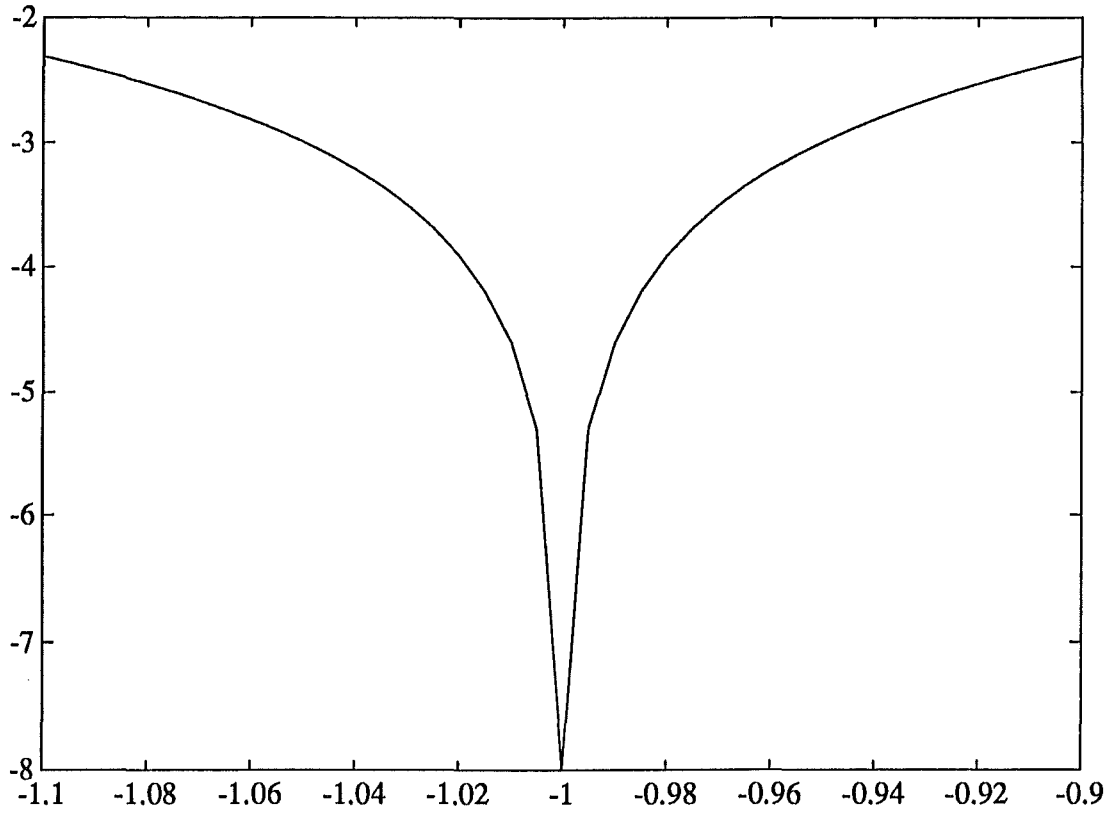


Figure 36

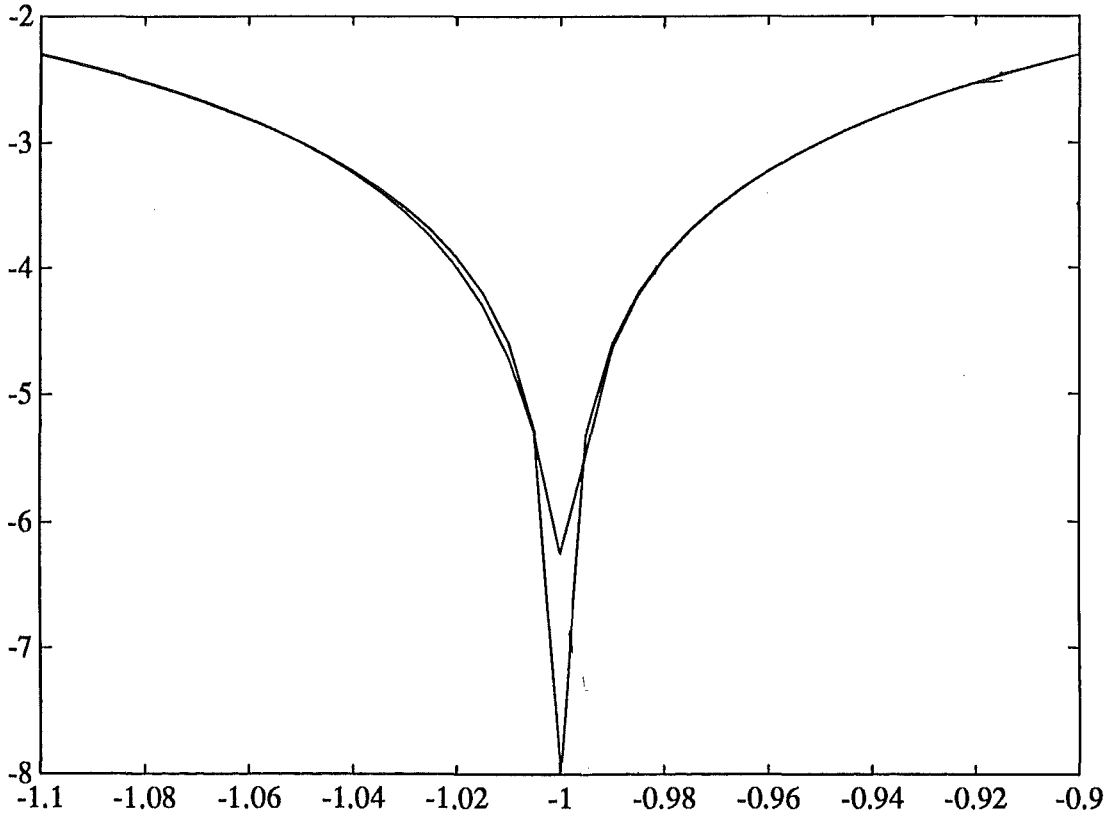


Figure 37

$R \subseteq \{x + \mathrm{i}y \in \mathbf{C} : |x| \leq 2, |y| \leq 2\}$ has now been defined so that $y(z)$ has a unique analytic continuation on R. It remains to calculate $y(z)$ at points in R. Clearly it is not sufficient to just write

$$y(z) = \frac{-a_1(z) + x\sqrt{d(z)}}{2a_2(z)}$$

and the attempt to do so results in imag($y(z)$) behaving as in Fig.38.

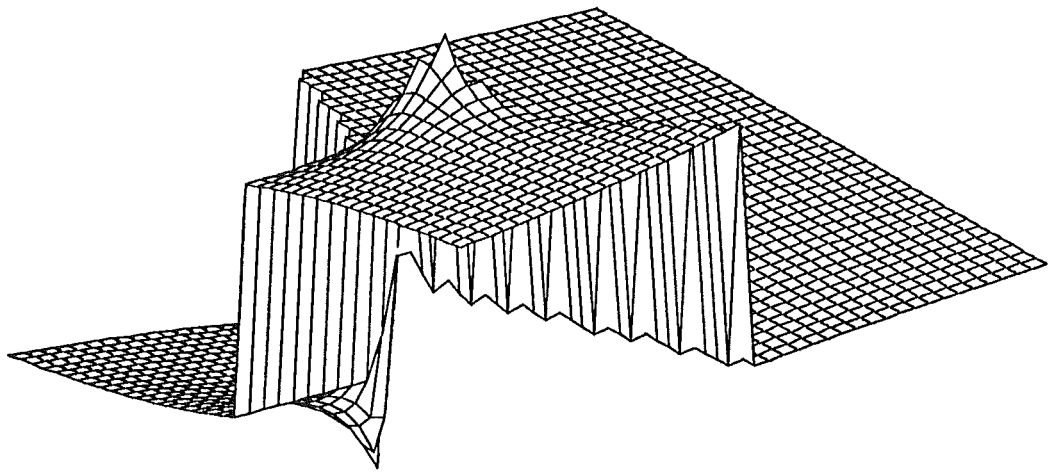


Figure 38

It is necessary to follow a path inside R "analytically" from the origin to each point. The following algorithm is used for this "analytic" procedure.

Algorithm.

The process is given for the first quadrant only. The rest are similar. (Note that $\mathbf{i}^2 = -1$).

sign $= 1$

$dz(0,0) = \sqrt{d(0)}$

arg1 $=$ arg$(dz(0,0))$

For $j = 0.1$ step 0.1 to 2 do

  $dz(0,j) =$ sign $\sqrt{d(j\mathbf{i})}$

  arg2 $=$ arg $(dz(0,j))$

  If $|\text{arg1} - \text{arg2}| > \frac{2\pi}{3}$ then

    sign $=-$ sign

    $dz(0,j) = -dz(0,j)$

    arg2 $=$ arg $(dz(0,j))$

  end

  arg1 $=$ arg2

end.

Note that $dz(0,j)$ now has the appropriate value of $\sqrt{d(z)}$ for $z = 0 + j\mathbf{i}$. To cover the remainder of points we step outward in lines from the imaginary axis as follows :

For $j = 0$ step 0.1 to 2 do

  sign $= 1$

  $z = j\mathbf{i}$

  arg1 $=$ arg$(dz(0,j))$

  For $k = 0.1$ step 0.1 to 2 do

    $z = k + j\mathbf{i}$

    $dz(k,j) =$ sign $\sqrt{d(z)}$

    arg2 $=$ arg $(dz(k,j))$

    If $|\text{arg1} - \text{arg2}| > \frac{2\pi}{3}$ then

      sign $=-$ sign

      $dz(k,j) = -dz(k,j)$

      arg2 $=$ arg $(dz(k,j))$

    end

    arg1 $=$ arg2

  end

end

It now remains only to form the matrix

$$\frac{-a_1(z) + zdz(k,j)}{2a_2(z)}, \; z = k + j\mathbf{i}$$

to get $y(z)$ on all points of the mesh. Graphs of the real and imaginary parts of $y(z)$ and $e(z) = y(z) - \log(1+z)$ along with contour maps of real($e(z)$) and imag($e(z)$) are now presented.

Fig	Real/Imag	Truncation
39	real(y(z))	$\pm\infty$
40	imag(y(z))	$\pm\infty$
41	real(e(z))	$\pm 10^{-1}$
42	imag(e(z))	$\pm 10^{-1}$
43	real(e(z))	$\pm 10^{-3}$
44	imag(e(z))	$\pm 10^{-3}$
45	real(e(z))	$\pm 10^{-5}$
46	imag(e(z)	$\pm 10^{-5}$
47	real(e(z))	$\pm 10^{-8}$
48	imag(e(z))	$\pm 10^{-8}$

Figures 49 and 50 are contour maps of real($e(z)$) and imag($e(z)$) with contours drawn at $\{\pm 10^{-3} \pm 10^{-4}, \ldots, \pm 10^{-8}\}$.

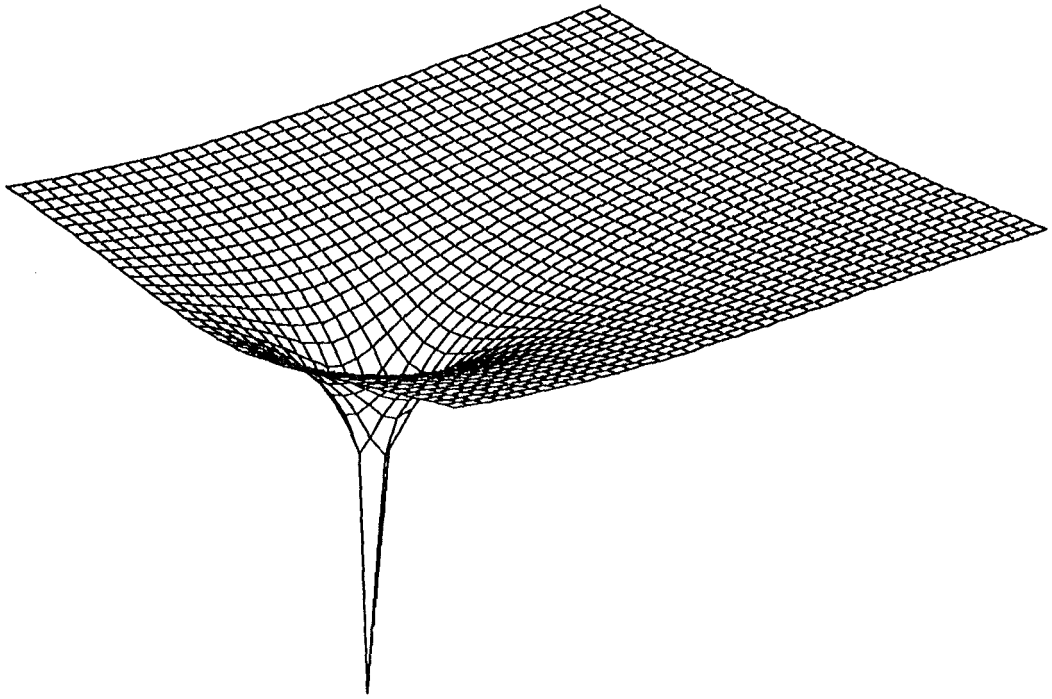


Figure 39 Real(y(z)). Truncation $\pm\infty$

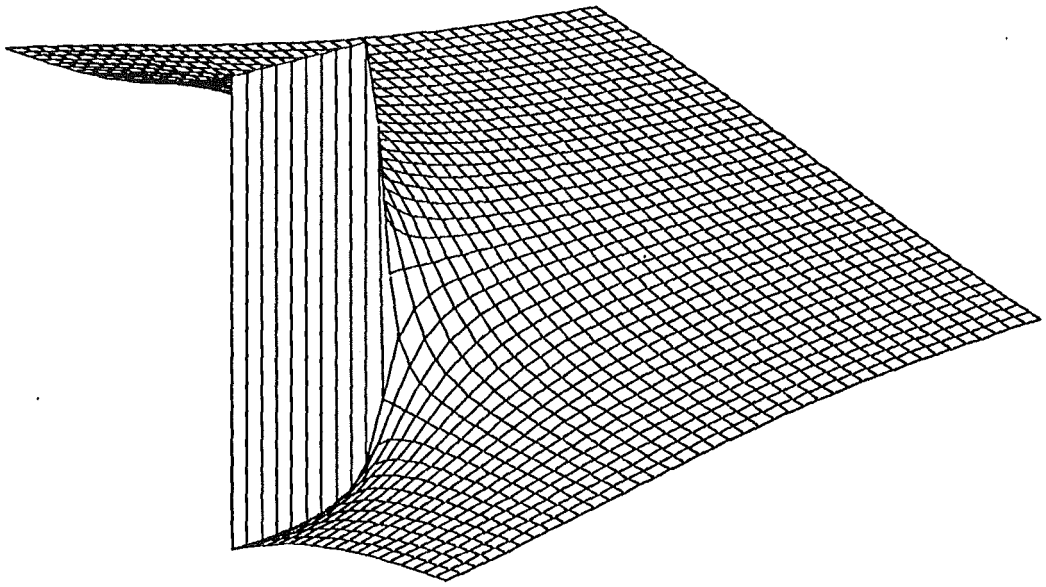


Figure 40 Imag(y(z)). Truncation $\pm\infty$

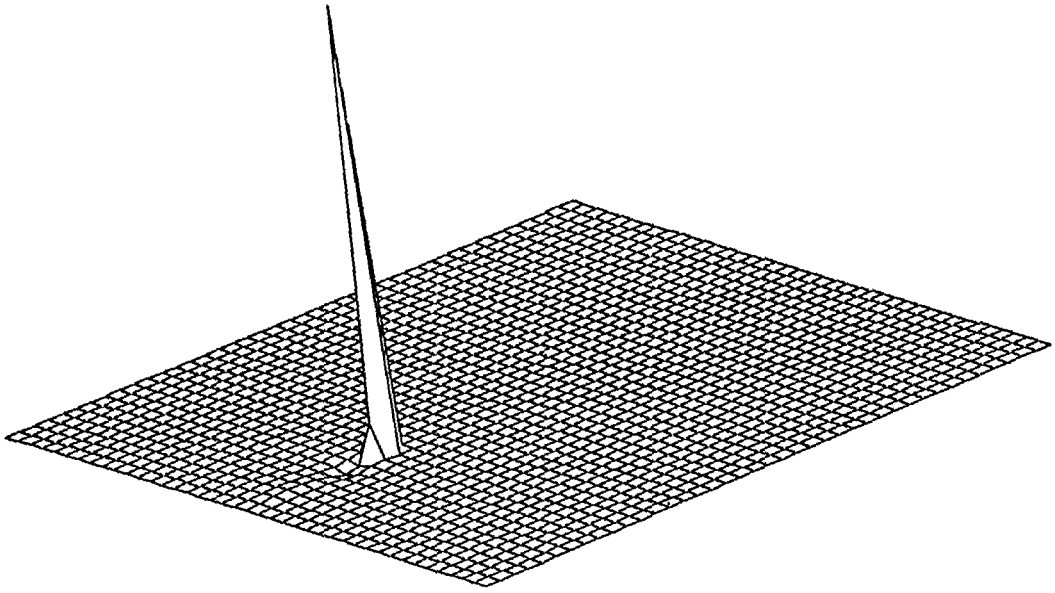


Figure 41 Real(e(z)). Truncation $\pm 10^{-1}$

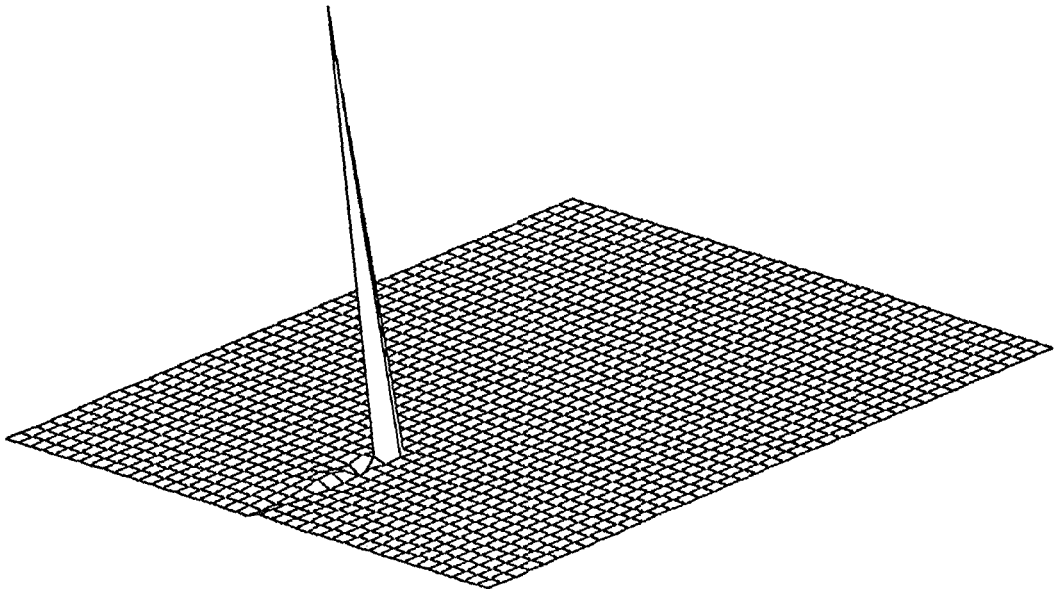


Figure 42 Imag(e(z)). Truncation $\pm 10^{-1}$

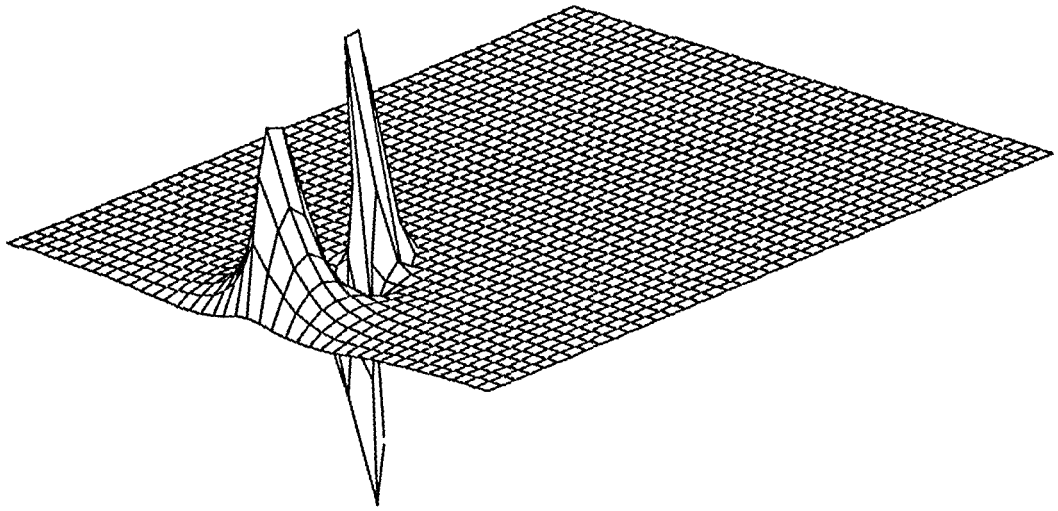


Figure 43 Real(e(z)). Truncation $\pm 10^{-3}$

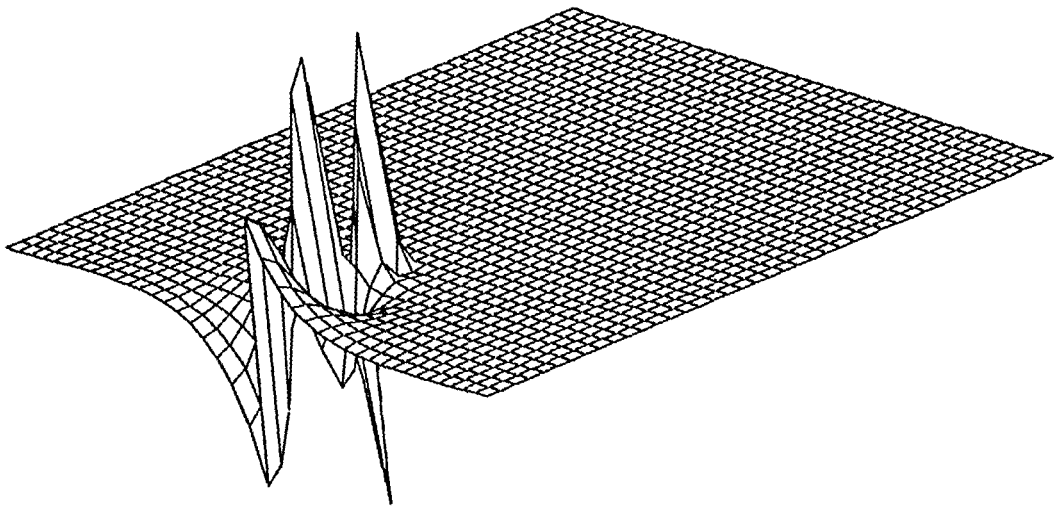


Figure 44 Imag(e(z)). Truncation $\pm 10^{-3}$

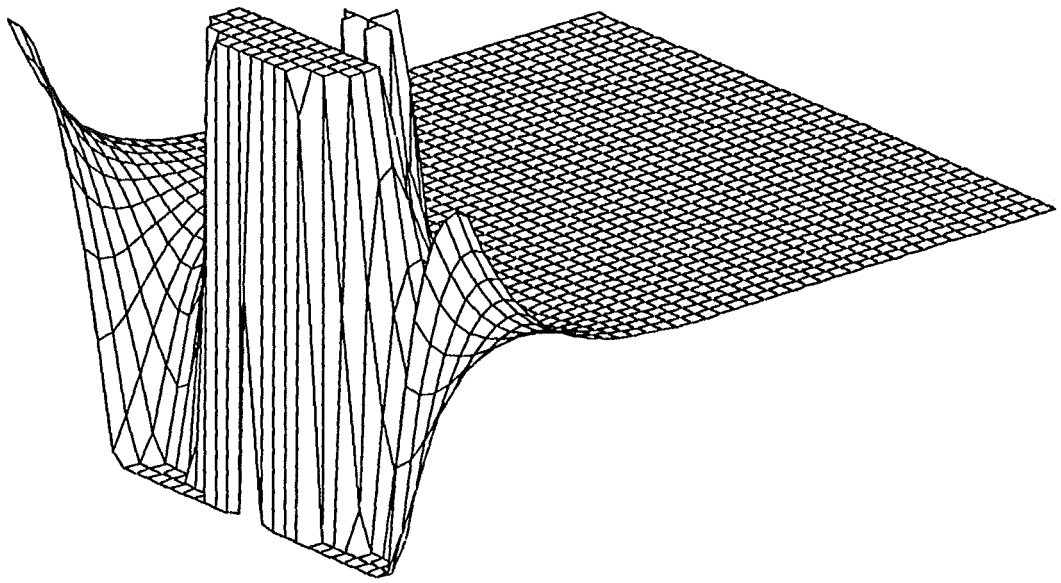


Figure 45 Real(e(z)). Truncation $\pm 10^{-5}$

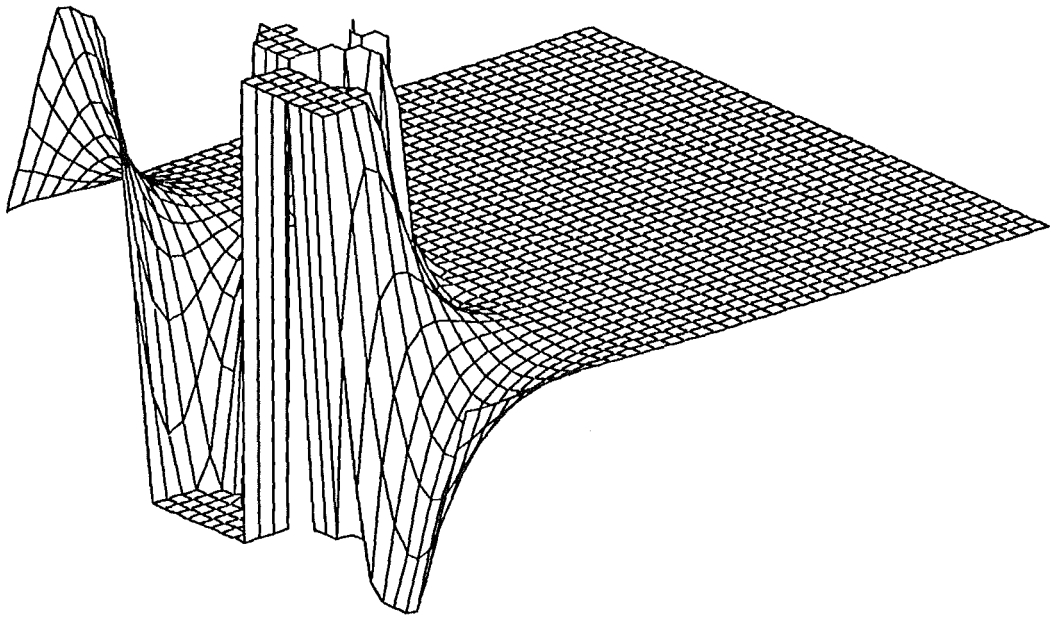


Figure 46 Imag(e(z)). Truncation $\pm 10^{-5}$

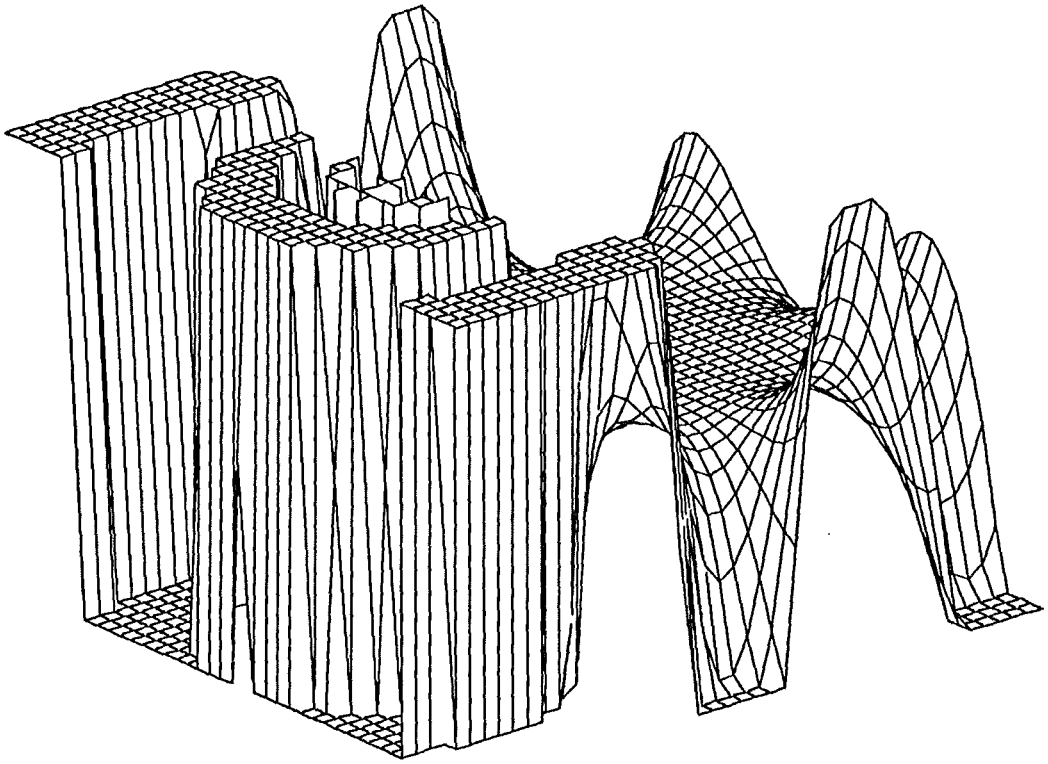


Figure 47 Real(e(z)). Truncation $\pm 10^{-8}$

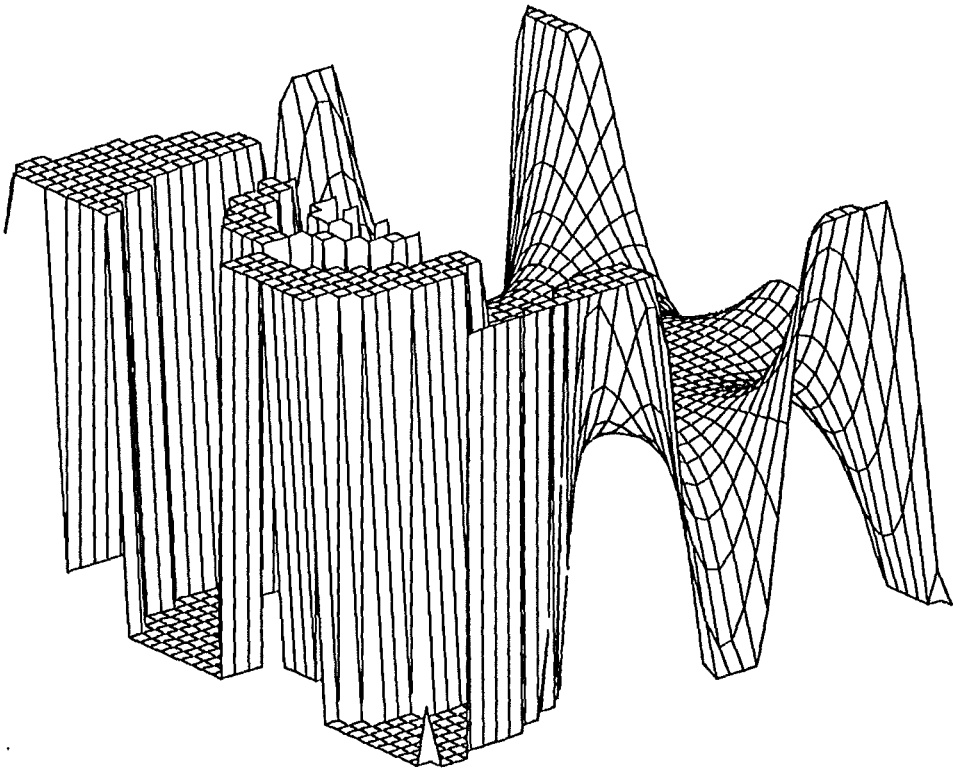


Figure 48 Imag(e(z)). Truncation $\pm 10^{-8}$

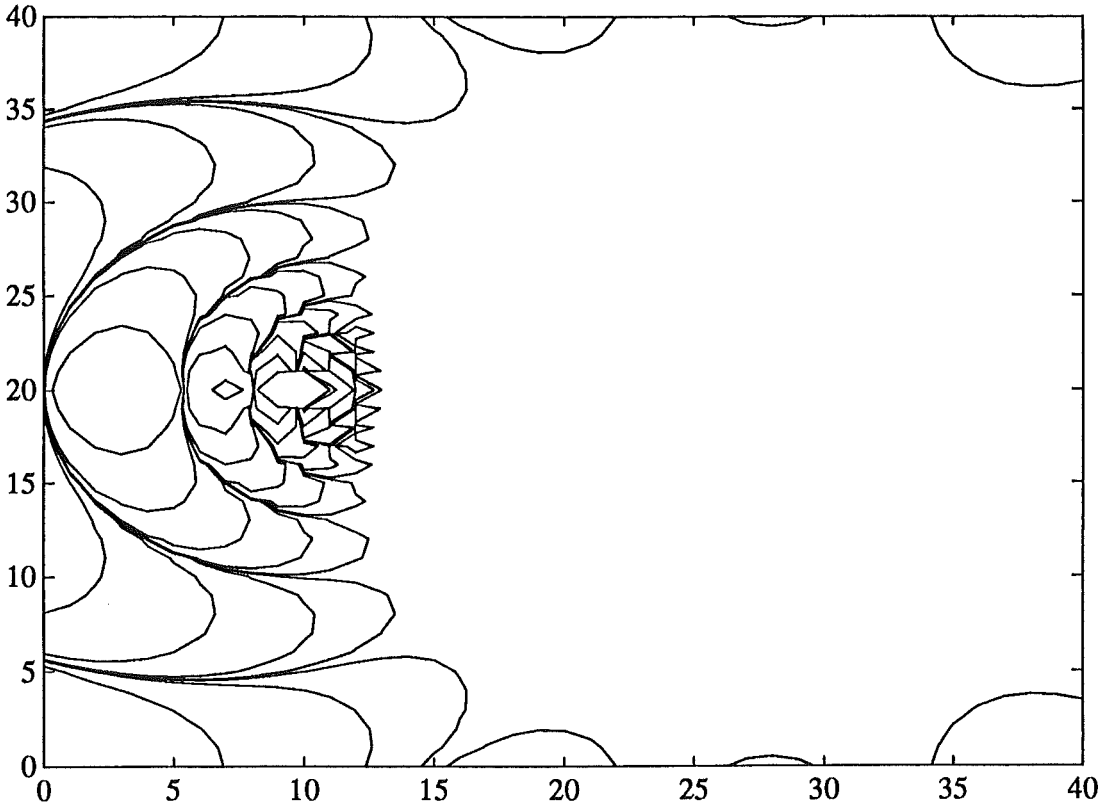


Figure 49 Contour map of Real(e(z)).

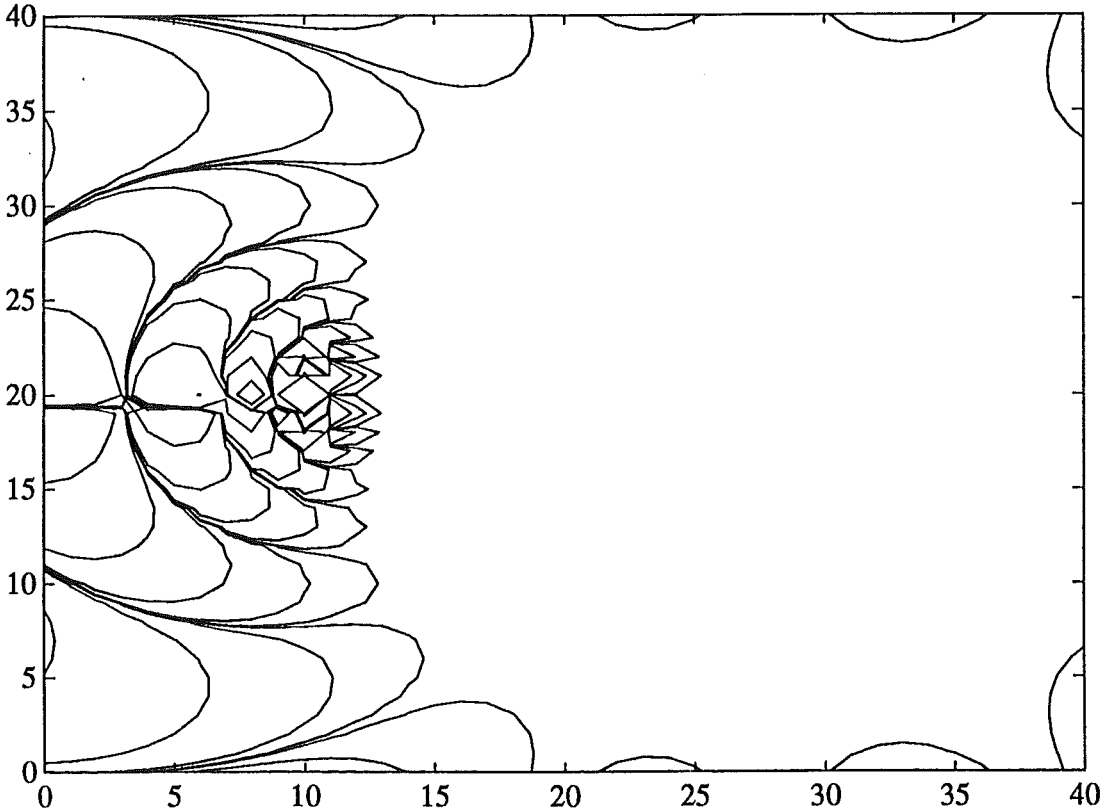


Figure 50 Contour map of Imag(e(z)).

A comparison with the (6,6) Padé approximation to log(1+x). Note that

$$p(x) = \frac{49x^6 + 1218x^5 + 7980x^4 + 20720x^3 + 23100x^2 + 9240x}{10x^6 + 420x^5 + 4200x^4 + 16800x^3 + 31500x^2 + 27720x + 9240}$$

and that

$$y(x) = f(x) + O\left(x^{13}\right)$$
$$p(x) = f(x) + O\left(x^{13}\right) .$$

Figures 51 and 52 are graphs of real($p(z)$) and imag($p(z)$) while figures 53 and 54 are contour maps of real($e(z)$) and imag($e(z)$) with contours drawn at $\{\pm 10^{-3}, \pm 10^{-4}, \ldots, \pm 10^{-8}\}$.

Clearly $p(x)$ is inferior to $y(x)$ as an approximation and examination of the Taylor polynomial of degree 12 shows that it is still less accurate.

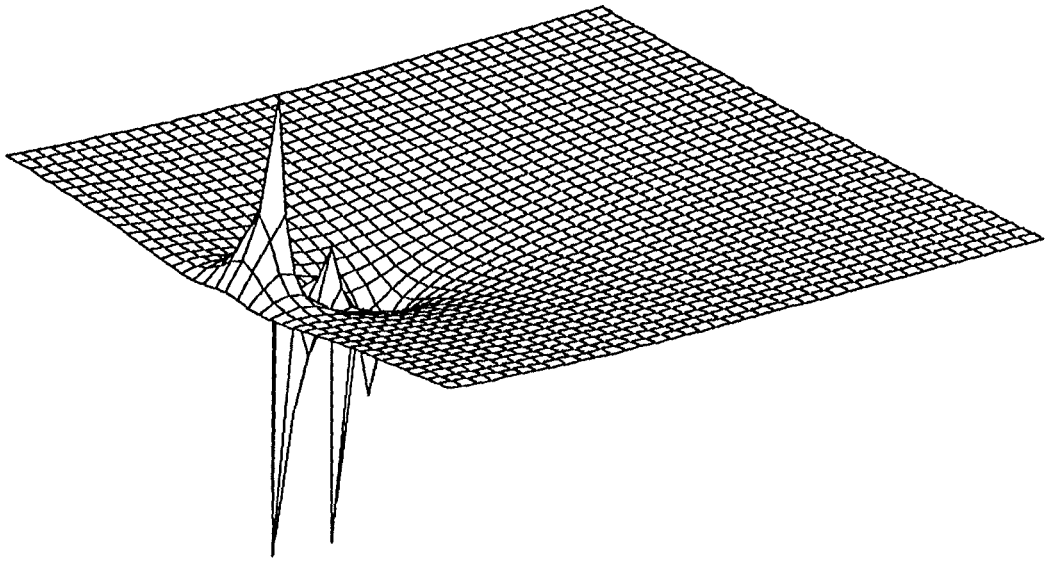


Figure 51 Real(p(z)).

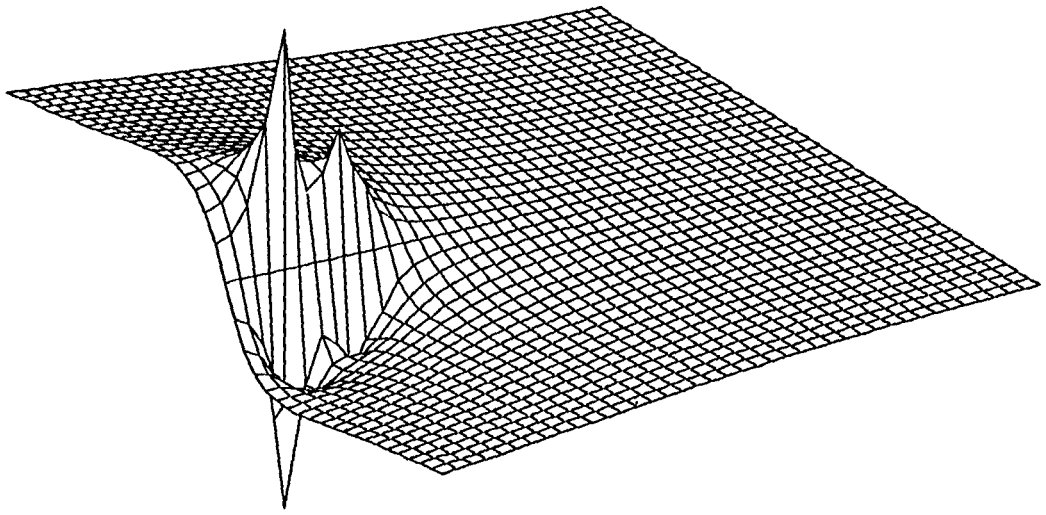


Figure 52 Imag(p(z)).

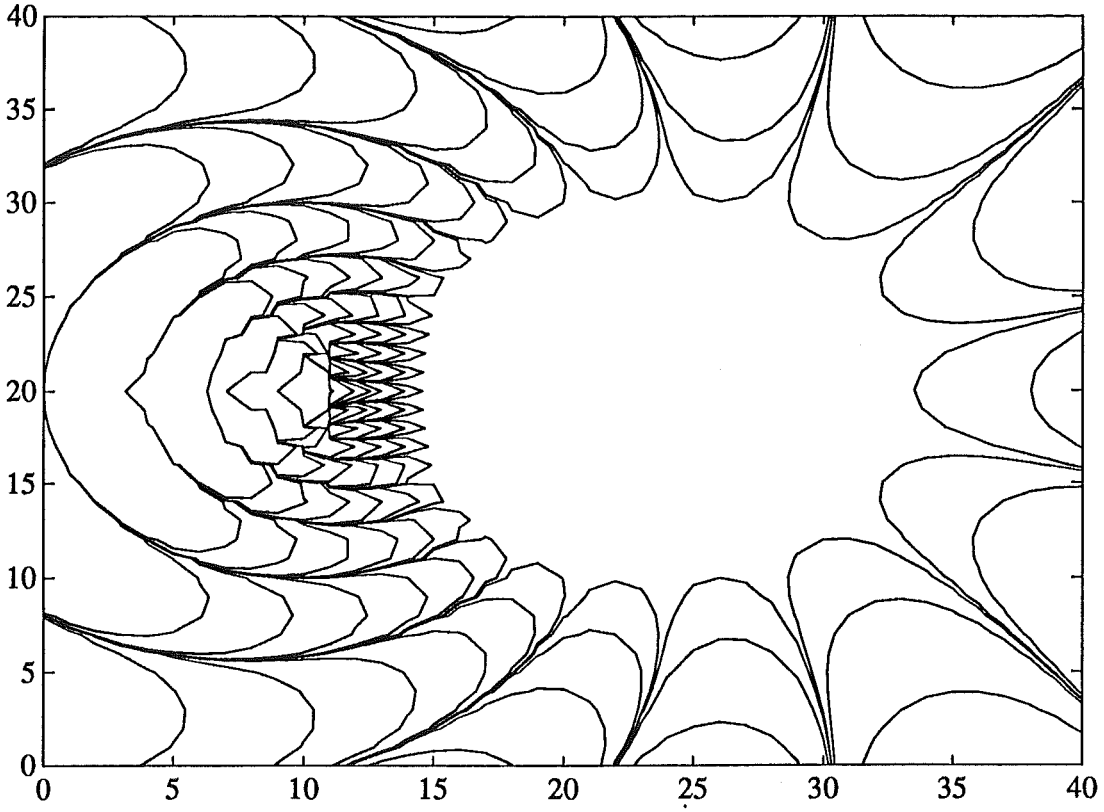


Figure 53 Contour map of Real(e(z)).

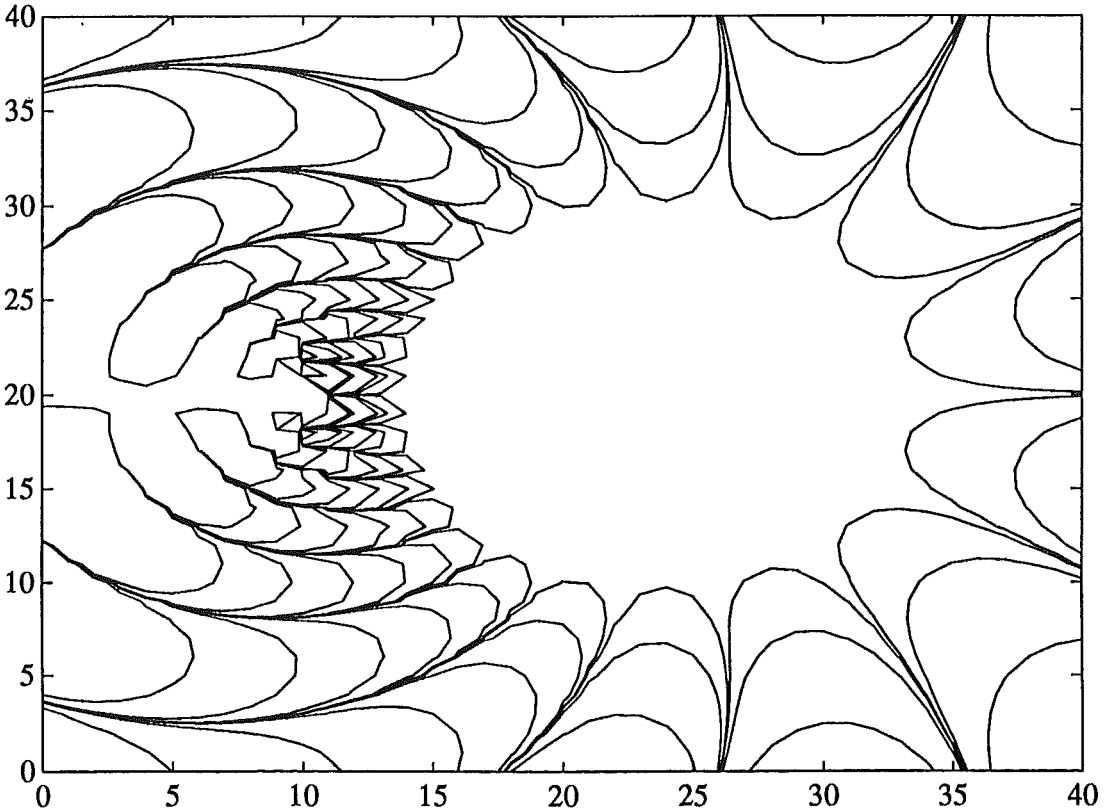


Figure 54 Contour map of Imag(e(z)).

3.3 Example 3

The (2,2,2) approximation to e^{-x} . Note that

$$\left(x^2+9x+24\right) f\left(x\right)^2-\left(8x^2+48\right) f\left(x\right)+x^2-9x+24=O\left(x^8\right)$$

so that

$$y\left(x\right)=\frac{8x^2+48-x\sqrt{60x^2+900}}{2\left(x^2+9x+24\right)}=f\left(x\right)+O\left(x^7\right) .$$

Figures 55 and 56 are graphs of real($y(z)$) and imag($y(z)$) while figures 57 and 58 are contour maps of real($e(z)$) and imag($e(z)$) with contours drawn at $\{\pm 10^{-3}, \pm 10^{-4}, \ldots, \pm 10^{-8}\}$

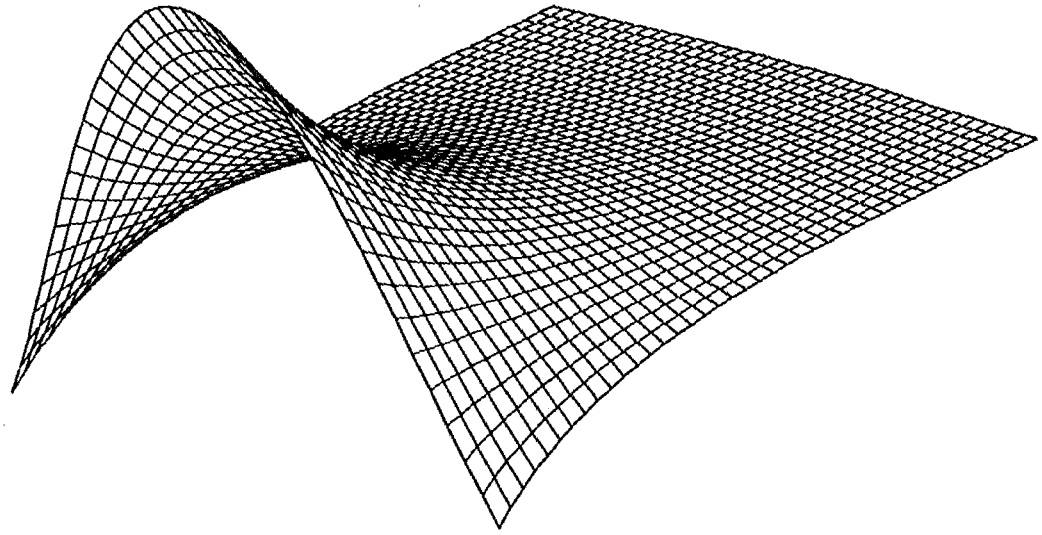


Figure 55 Real(y(z)).

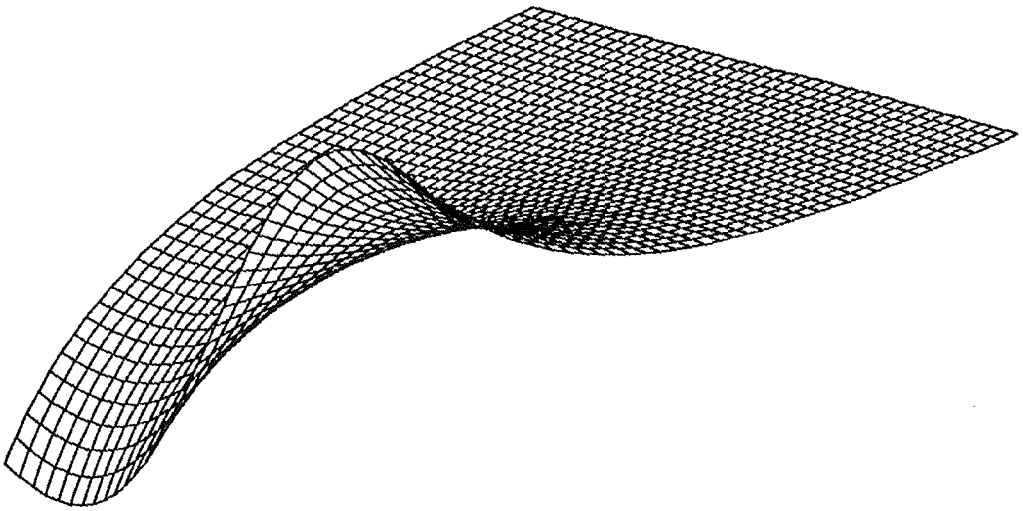


Figure 56 Imag(y(z)).

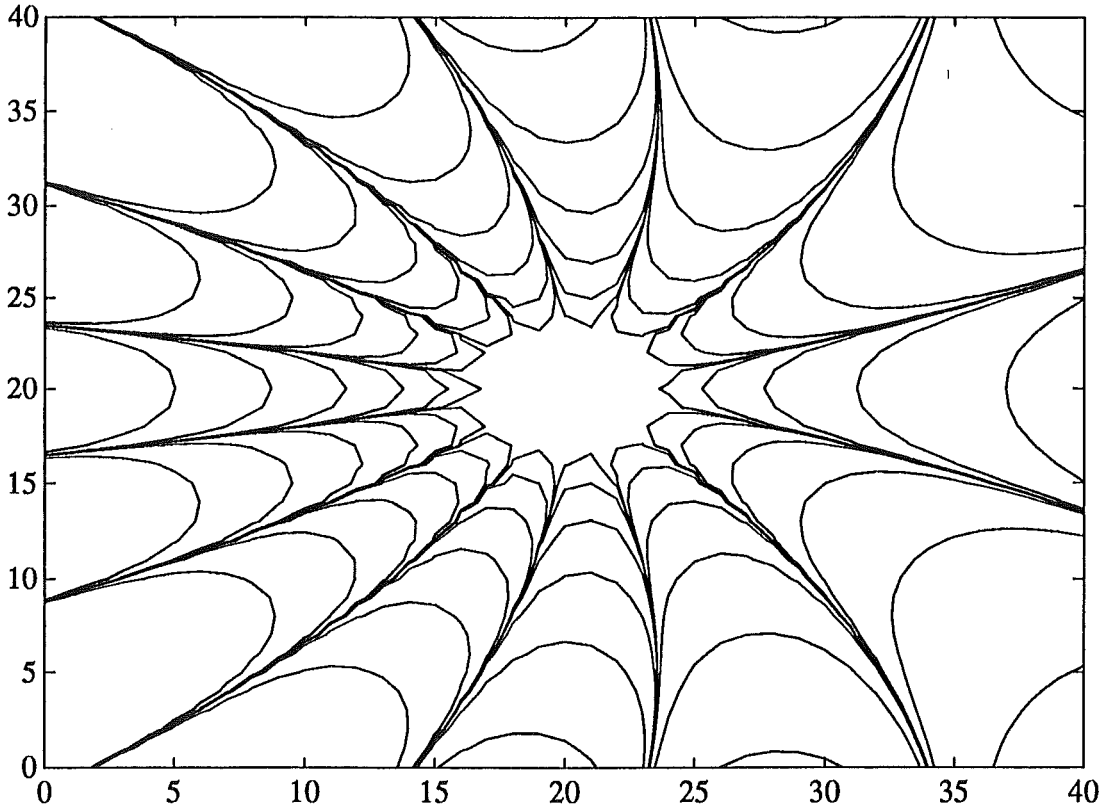


Figure 57 Contour map of Real(e(z)).

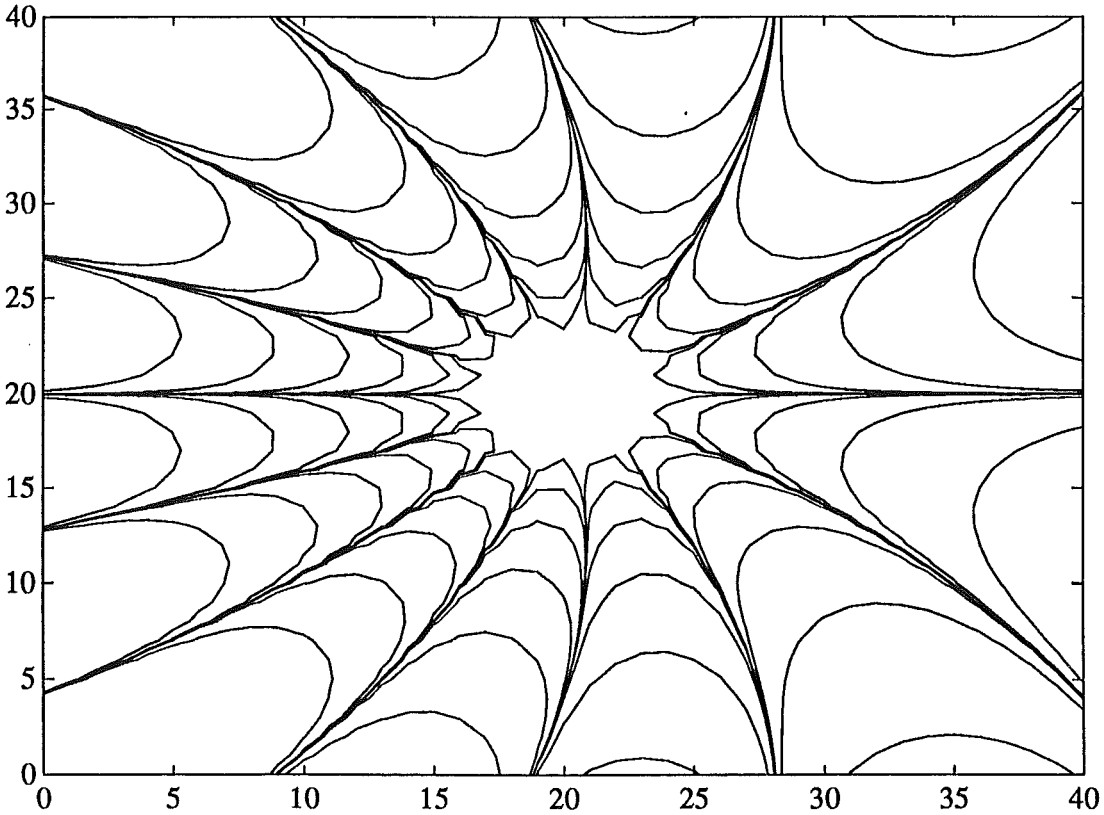


Figure 58 Contour map of Imag(e(z)).

A comparison with the Taylor polynomial of degree 6 . The Taylor polynomial approximation to e^{-x}, of degree 6 is

$$t(x) = 1 - x + \frac{x^2}{2!} - \ldots + \frac{x^6}{6!} = f(x) + O\left(x^7\right) .$$

Figures 59 and 60 are the usual contour maps of real$(e(z))$ and imag$(e(z))$. Certainly $t(x)$ is inferior to $y(x)$ but the difference is markedly less dramatic than with $\log(1+x)$. This is to be expected since e^{-x} is a smoother function, with a faster converging power series than $\log(1+x)$.

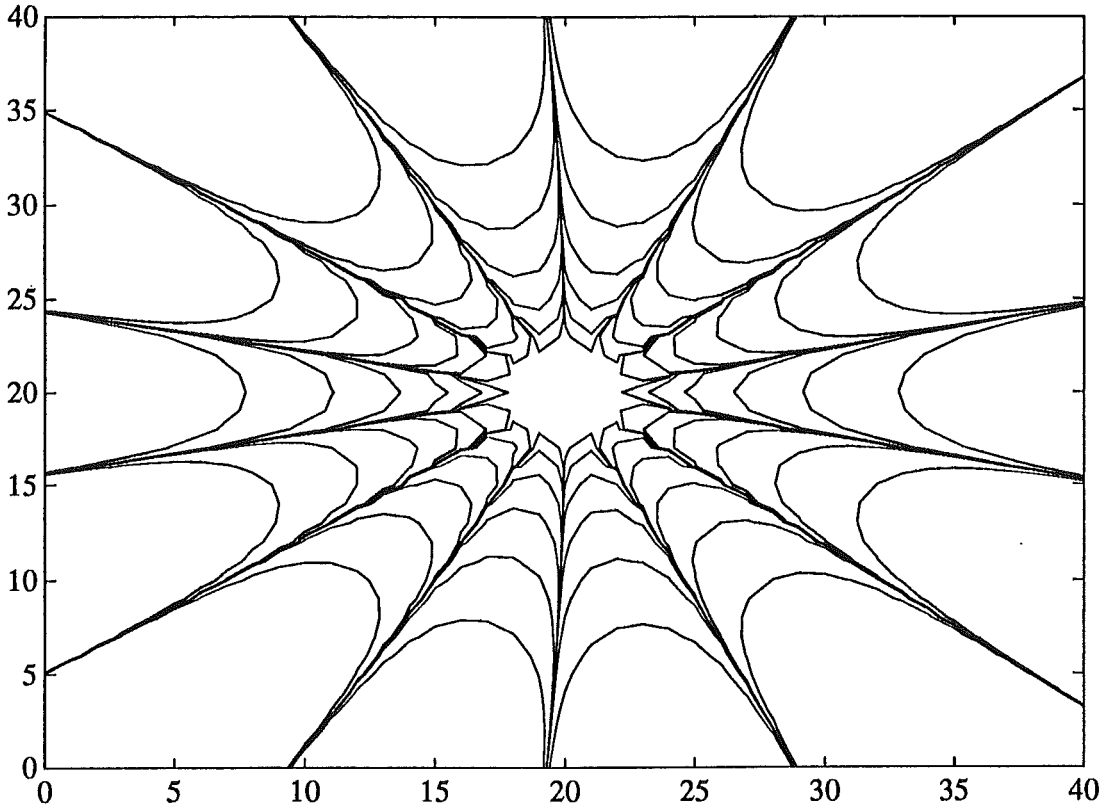


Figure 59 Contour map of Real(e(z)).

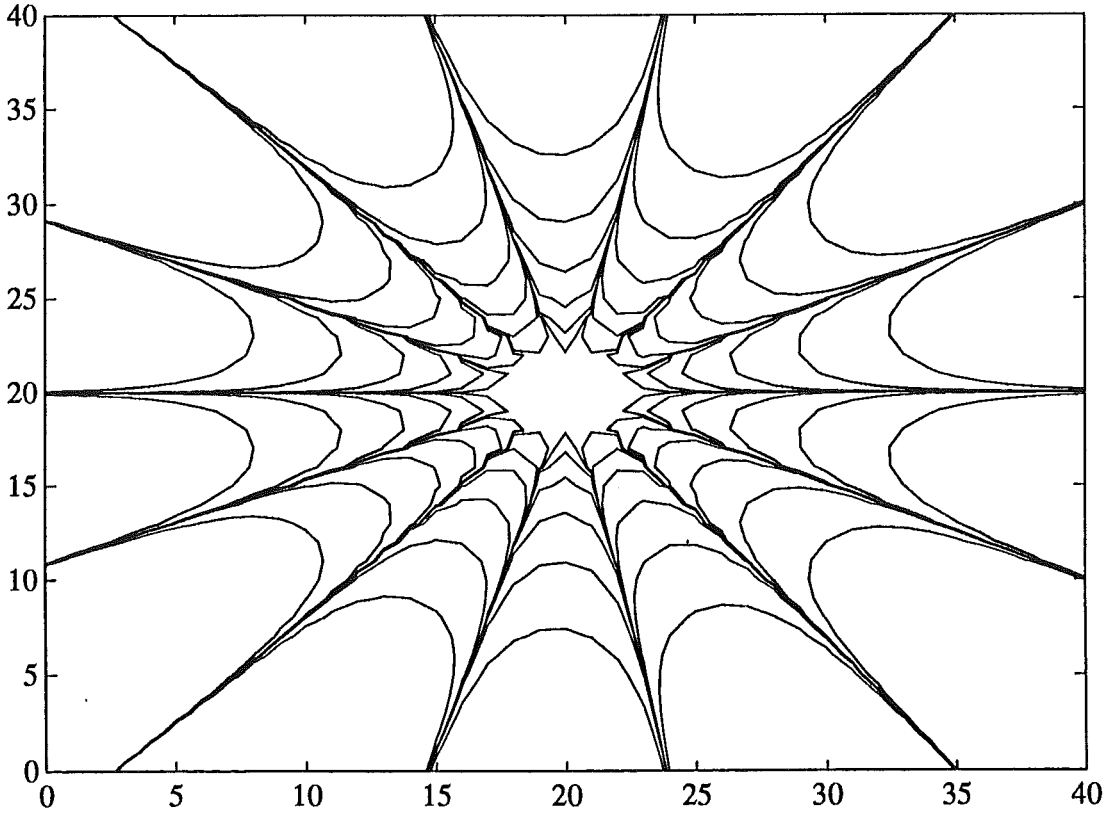


Figure 60 Contour map of Imag(e(z)).

3.4 Example 4

The (5,5,5) approximation to e^{-x} . Note that

$$\begin{aligned}&\left(x^5+45x^4+885x^3+9450x^2+54495x+135135\right)f(x)^2\\&+\left(64x^5+7680x^3+161280x\right)f(x)\\&+\left(x^5-45x^4+885x^3-9450x^2+54495x-135135\right)=O\left(x^{17}\right)\end{aligned}$$

so that

$$y(x)=\frac{-64x^5-7680x^3-161280x+\sqrt{D(x)}}{2\left(x^5+45x^4+885x^3+9450x^2+54495x+135135\right)}=f(x)+O\left(x^{17}\right).$$

where

$$D(x)=4092x^{10}+984060x^8+79459380x^6+2497294800x^4+24348624300x^2+73045872900$$

Figures 61 and 62 are contour maps of real($e(z)$) and imag($e(z)$) this time with contours drawn at $\{\pm10^{-8},\pm10^{-9},\ldots,\pm10^{-11}\}$.

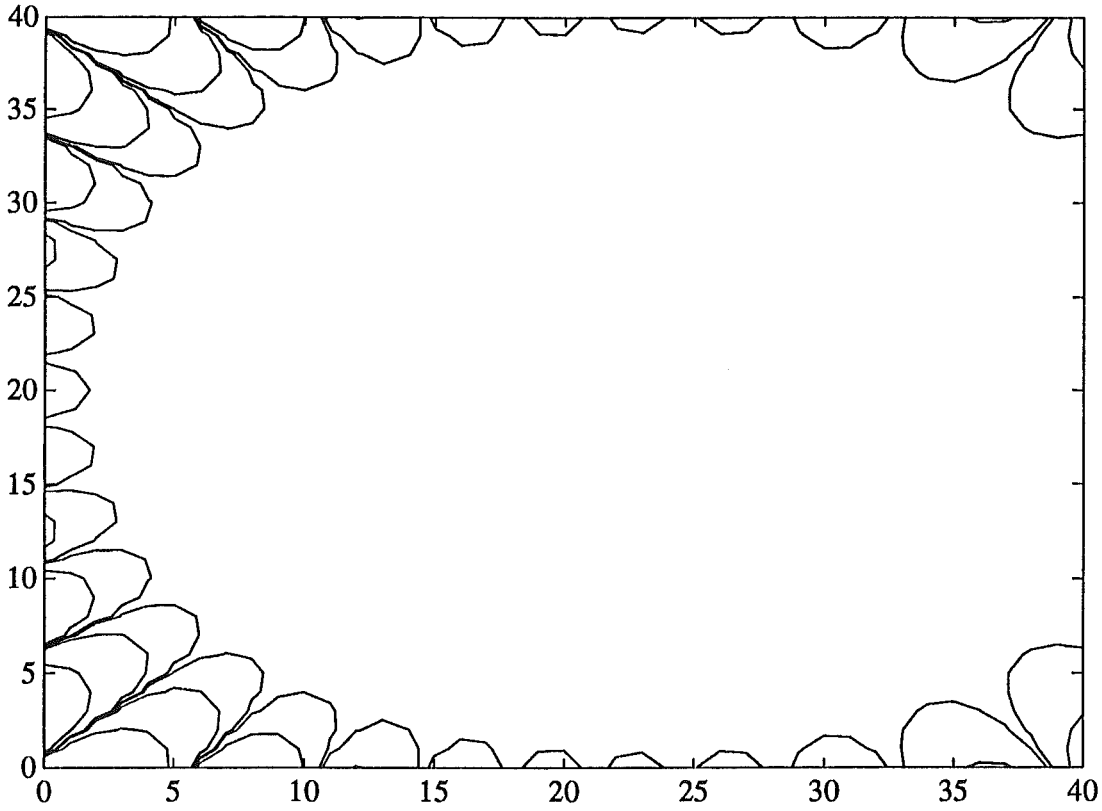


Figure 61 Contour map of Real(e(z)).

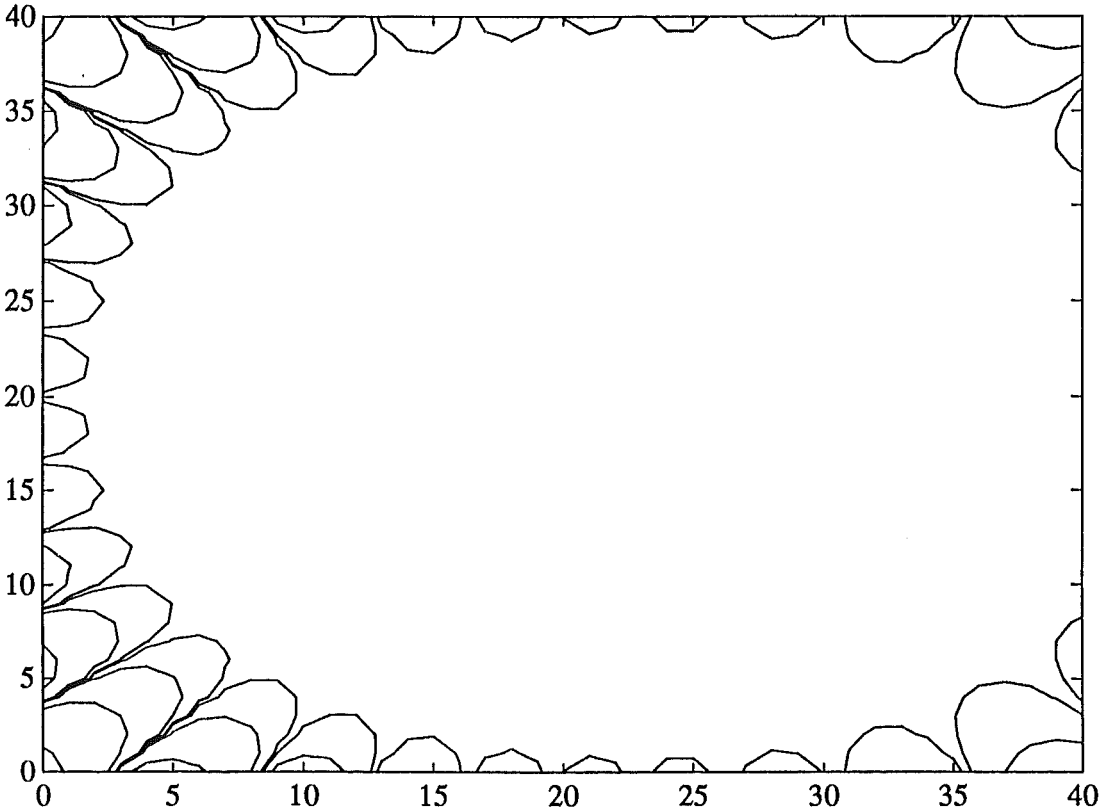


Figure 62 Contour map of Imag(e(z)).

A comparison with the Taylor polynomial of degree 16 . Note that $t(x) = f(x) + O(x^{17})$. Figures 63 and 64 are contour maps of real$(e(z))$ and imag$(e(z))$ with contours drawn at $\{\pm 10^{-8}, \ldots, \pm 10^{-11}\}$.

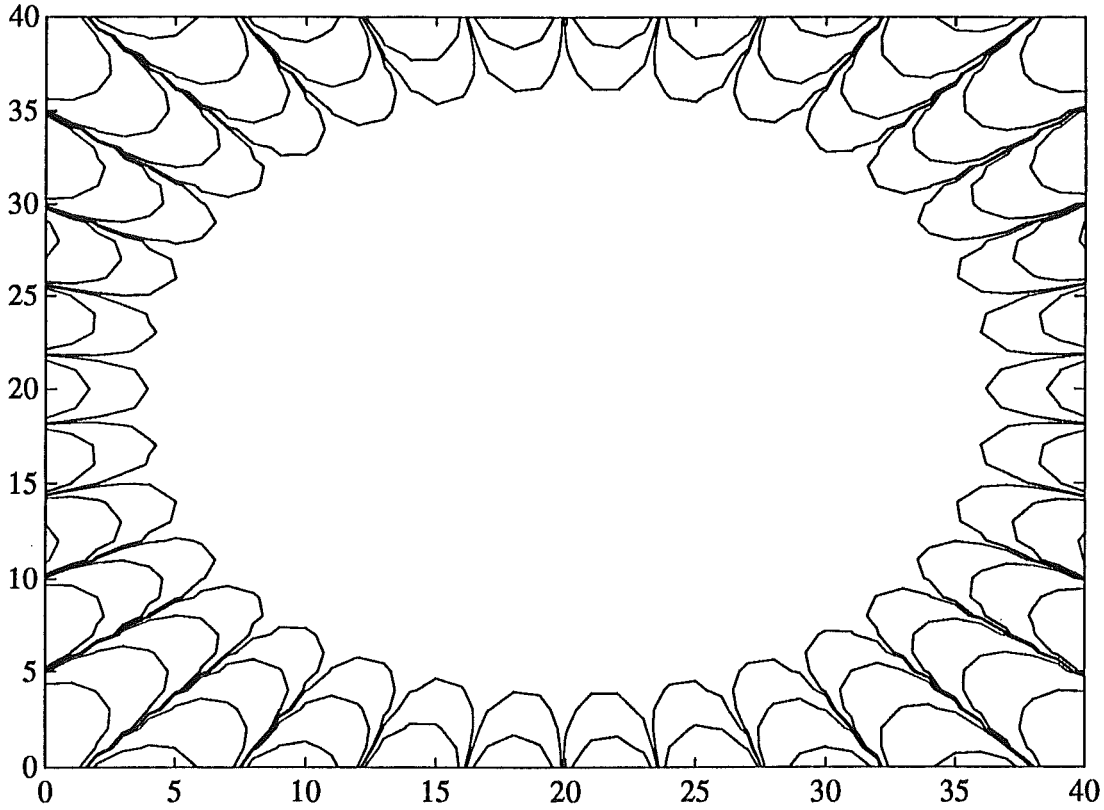


Figure 63 Contour map of Real(e(z)).

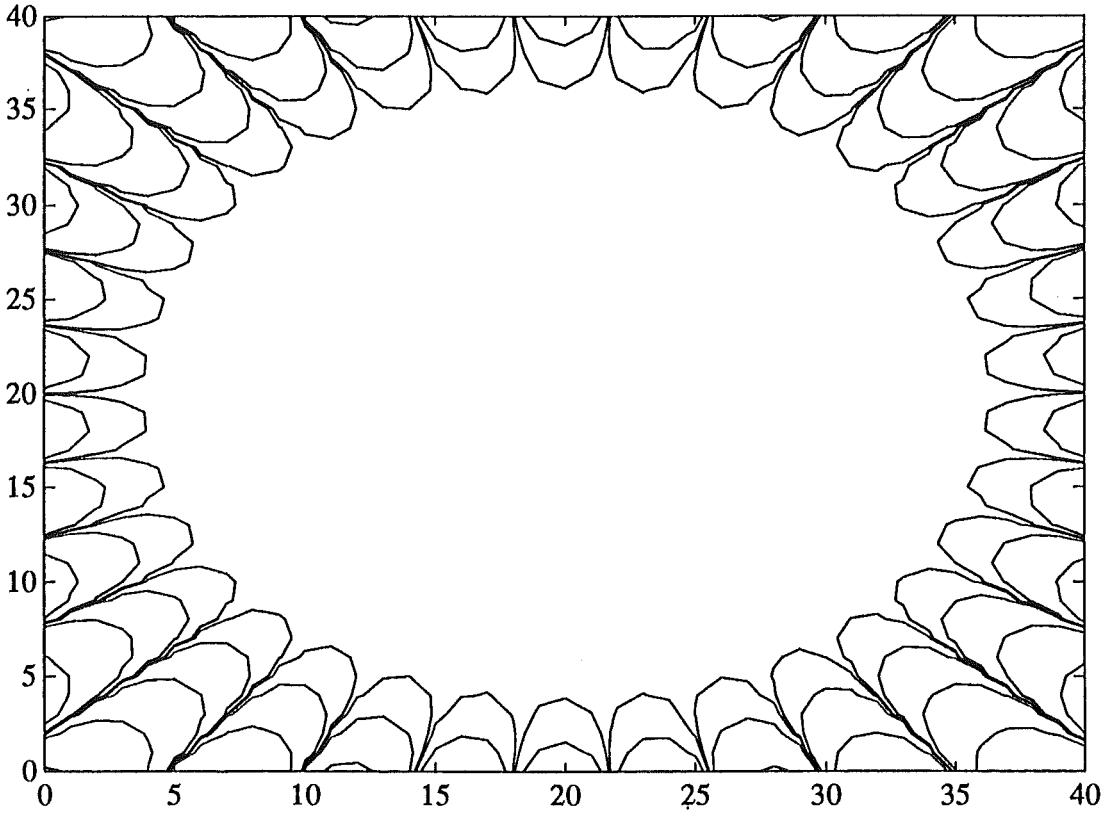


Figure 64 Contour map of Imag(e(z)).

4. Conclusion

This chapter has attempted to assess some qualitative properties of the quadratic approximation by means of detailed investigation of some particular examples. Particular attention was paid to the size of the region over which this was a good approximation and to comparisons of the accuracy of the quadratic approximation with the more traditional Padé and Taylor approximations. An algorithm for obtaining a smooth, analytic approximation over the wider region was given.

It is clear from these examples that when using the quadratic approximation, care must be taken in defining the region of the approximation, and particularly in the placement of cuts from the branch points. However if this is done in a sensible manner, the approximations in these examples appear to be significantly better than the usual approximations, particularly for a function with some branch point structure. It is also of interest to note that the approximation is able to accurately represent this branch point structure. It is apparent from the contour maps of the error function, that the quadratic approximation to the log function is at least two significant figures better than the corresponding Padé approximation.

The function $\exp(-x)$ was used as an example of a smooth function. The quadratic approximation to this function is also an improvement over the traditional approximations. However, the roughly one significant figure improvement in the error over the corresponding Taylor polynomial approximation does not appear to justify the additional computational cost for functions of this type. It is interesting to compare these results with the specific numerical results obtained by Borwein [1].

5. References

1. P.B. Borwein (1986) : *Quadratic Hermite-Padé Approximation to the Exponential Function.* Constr. Approx., **2** : 291–302.
2. J. H. Curtiss (1978) : Introduction to Functions of a Complex Variable. New York : Marcel Dekker Inc.
3. E. Hille (1962) : Analytic Function Theory, vol II. Boston : Ginn and Company.

CHAPTER 5

MORE QUALITATIVE RESULTS FOR THE QUADRATIC HERMITE-PADÉ APPROXIMATION

1. Introduction

In this chapter some of the characteristics of the quadratic Hermite-Padé approximation which were addressed in Chapter 4 are examined in further detail. In particular it is shown that in the examples studied it is possible to ignore many of the spurious singularities that occur in the approximation. This leads us to examine the approximation over a much larger region than was the case in Chapter 4. The examples chosen for study are $\cos(x)$, $\log(1+x)$ and $\sqrt[3]{1+x}$. The $(4,4,4)$ approximation is examined in all cases and comparisons made with the appropriate Padé approximations.

2. Examples

2.1 Example 1 : $\cos(x)$

The $(4,4,4)$ approximation to $\cos(x)$. Note that

(i)

$$\begin{aligned}&\left(11041x^4+953925x^2+30370095\right)f(x)^2\\&+\left(-1196192x^4-134400x^2-324253440\right)f(x)\\&+5459071x^4-132576150x^2+293883345=O\left(x^{16}\right)\end{aligned}$$

so that, using the results of Chapter 3, the approximation is

$$y(x)=\frac{-a_1(x)-\sqrt{D(x)}}{2a_2(x)}.$$

This Hermite-Padé form is, in fact, the best choice from the 2 dimensional space of (4,4,4) forms (see Chapter 6, §3, for further explanation) so is of greater order of accuracy than might otherwise be expected.

(ii)

$$D(x) = 1189780889220x^8 - 14653547716500x^6 + 605478537140400x^4 + 15071189726092500x^2 + 69439232925562500$$

so the roots of $D(x)$ are

$x = \pm 2.8245\mathbf{i}$

$x = \pm 2.8489\mathbf{i}$

$x = \pm(4.7026 \pm 2.8123\mathbf{i})$

Note also that the roots of $a_2(x)$ are $x = \pm 2.1503 \pm 6.9154\mathbf{i}$

In a manner similar to that of Chapter 4 Example 2 cuts are taken between the zeroes of small separation, namely $\{x\mathbf{i} : x \in (-2.8489, -2.8245)\}$ and $\{x\mathbf{i} : x \in (2.8245, 2.8489)\}$. Graphs and contour maps of the error function $e(z) = y(z) - \cos(z)$ drawn with PC-Matlab are now given. The region shown is $\{x + \mathbf{i}y \in \mathbf{C} : |x|, |y| \leq 5\}$ with a mesh spacing of 0.25 and $e(z)$ has been truncated at ± 1.

Figures 1 and 2 are graphs of real($e(z)$). Figures 3 and 4 are contour maps of real($e(z)$) and imag($e(z)$) with contours drawn at $\{\pm 1, \pm 10^{-1}, \pm 10^{-2}, \ldots, \pm 10^{-5}\}$.

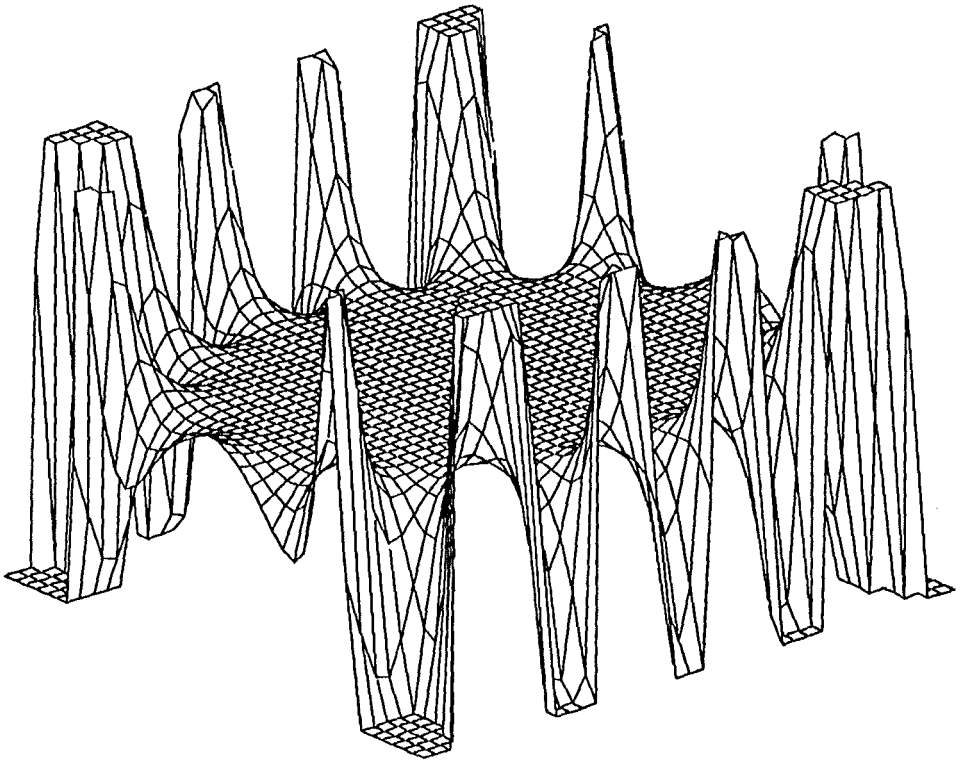


Figure 1 Real(e(z)). Truncation ± 1

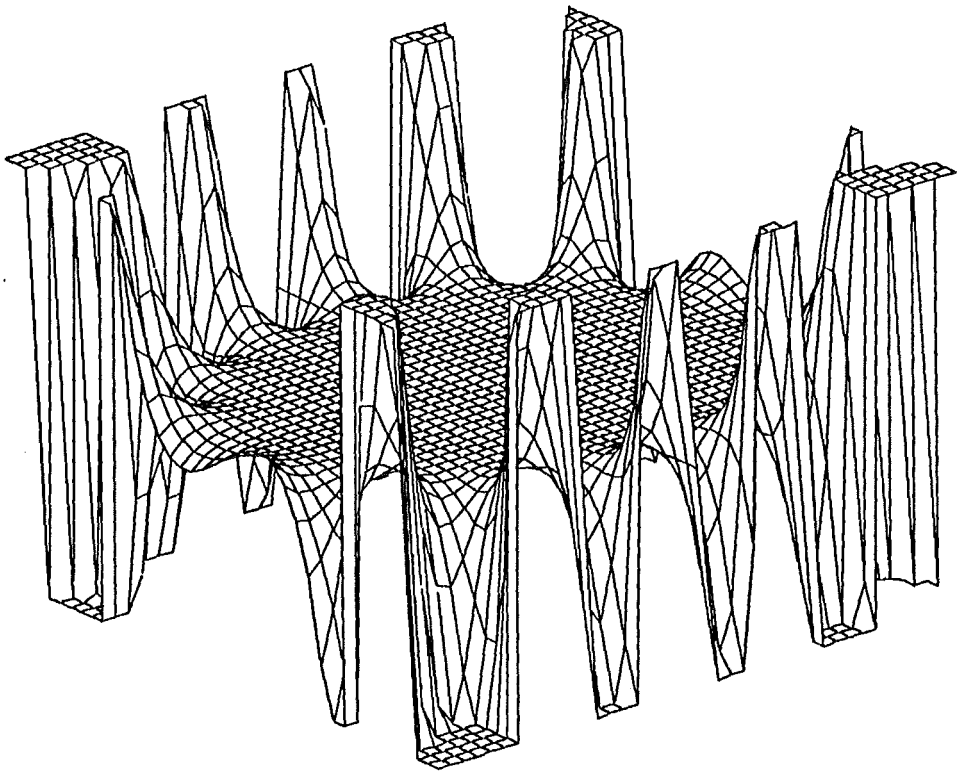


Figure 2 Imag(e(z)). Truncation ± 1

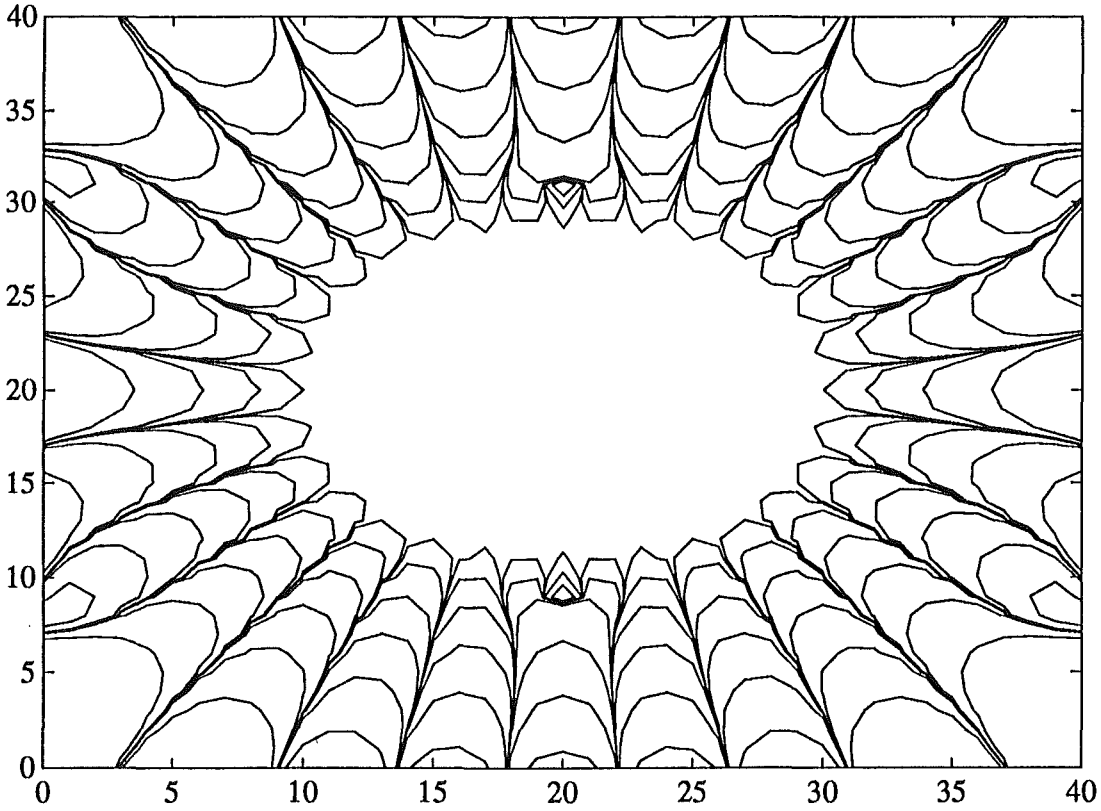


Figure 3 Contour map of Real(e(z)).

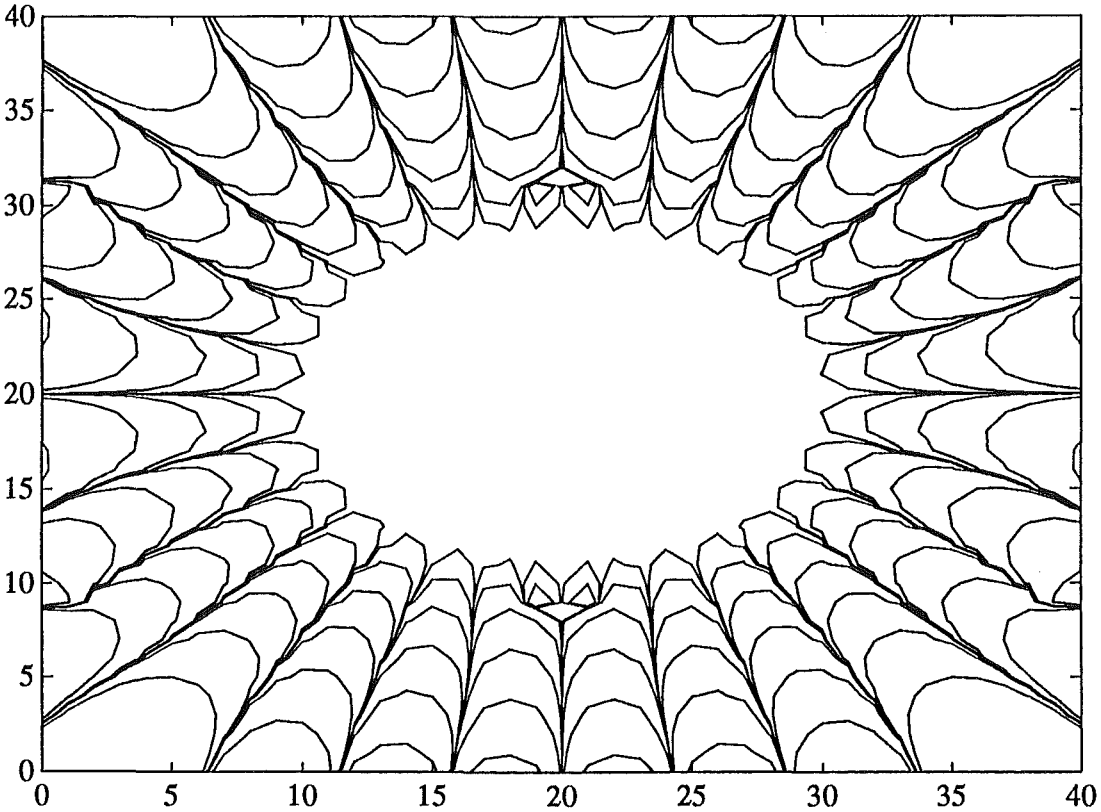


Figure 4 Contour map of Imag(e(z)).

The $(6,8)$ Padé approximation to $\cos(x)$. There is no $(7,7)$ Padé approximtion which matches $f(x)$ up to $O(x^{16})$ so the $(6,8)$ approximation has been chosen instead. Note that

$$p(x) = \frac{45469x^8 - 7029024x^6 + 348731040x^4 - 5269904640x^2 + 10983772800}{9336x^6 + 2064720x^4 + 221981760x^2 + 10983772800}$$

and that

$$y(x) = f(x) + O\left(x^{16}\right)$$
$$p(x) = f(x) + O\left(x^{16}\right) .$$

Figures 5 and 6 are graphs of real$(e(z))$ and imag$(e(z))$ (where $e(z) = p(z) - \cos(z)$) truncated at ± 1. Figures 7 and 8 are contour maps of real$(e(z))$ and imag$(e(z))$ with contours drawn at $\{\pm 1, \ldots, \pm 10^5\}$.

It is clear that $y(x)$ is much superior to $p(x)$ as an approximation to $\cos(x)$ over a considerably wider area.

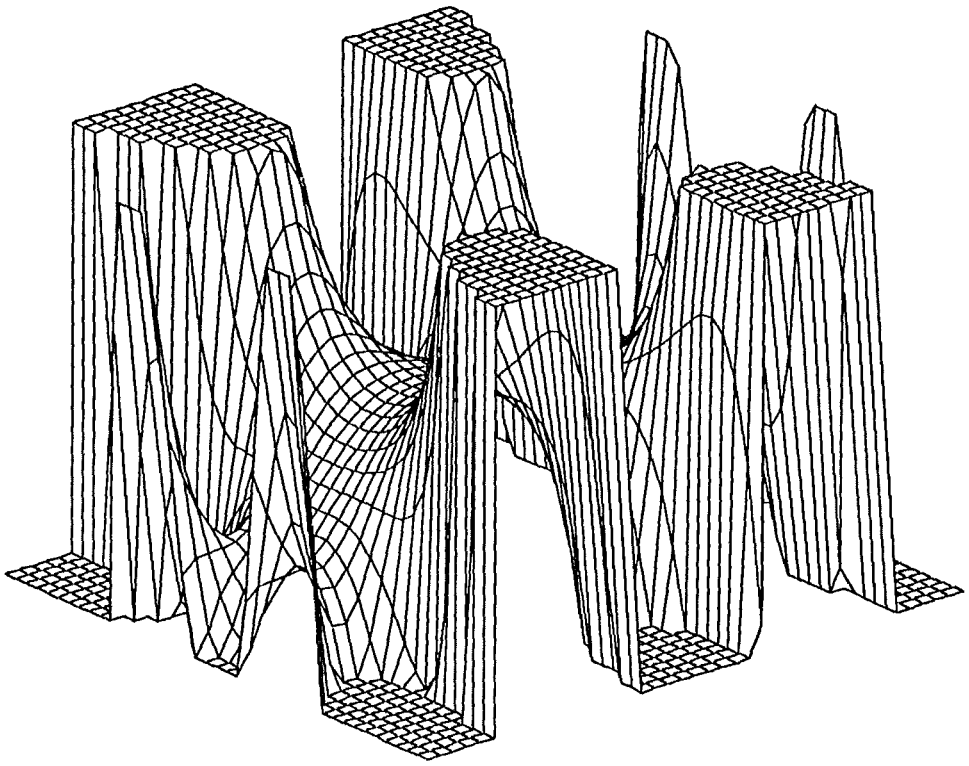


Figure 5 Real(e(z)). Truncation ± 1

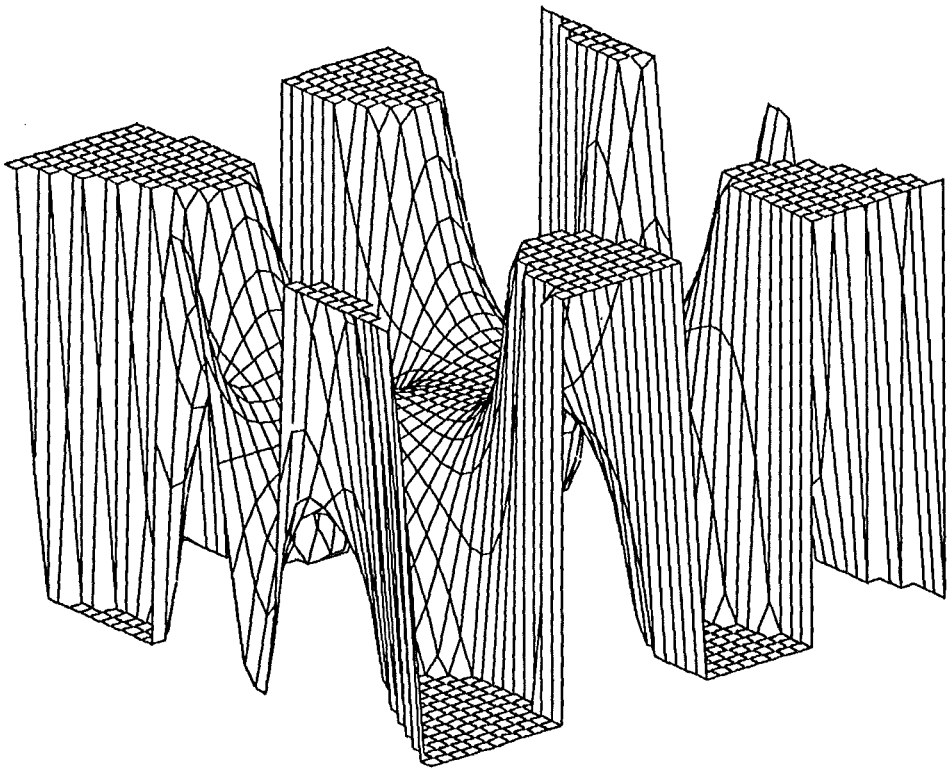


Figure 6 Imag(e(z)). Truncation ± 1

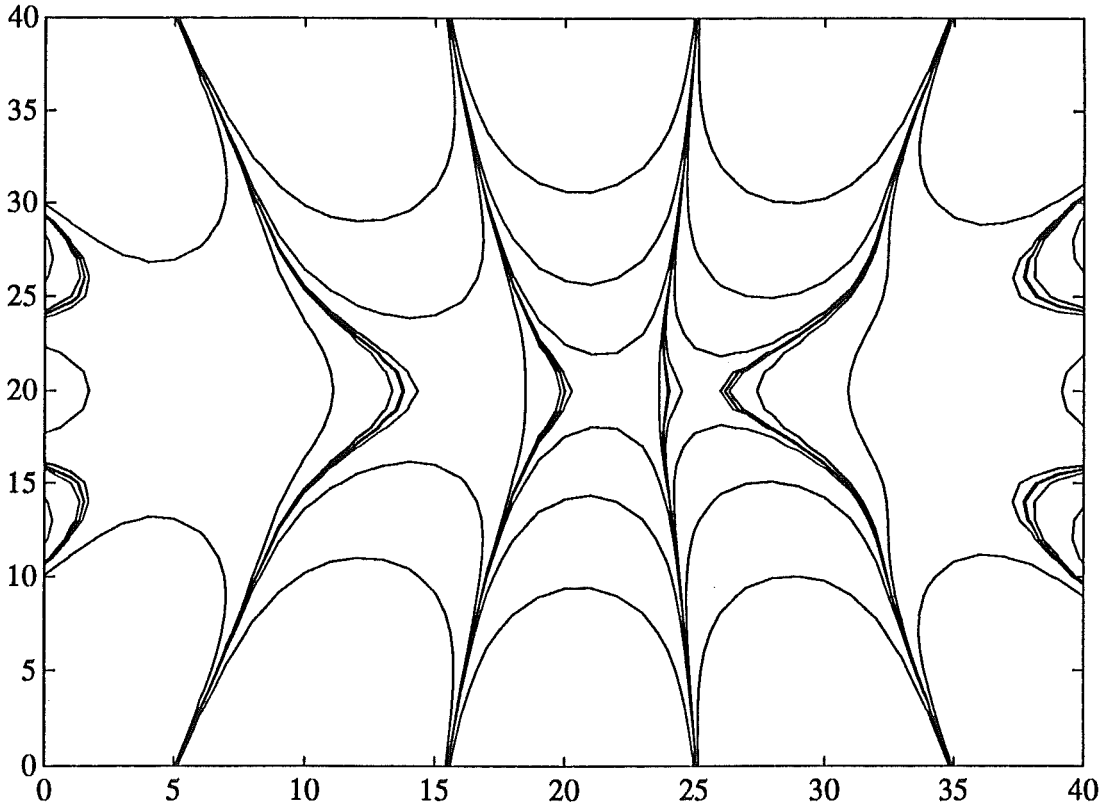


Figure 7 Contour map of Real(e(z)).

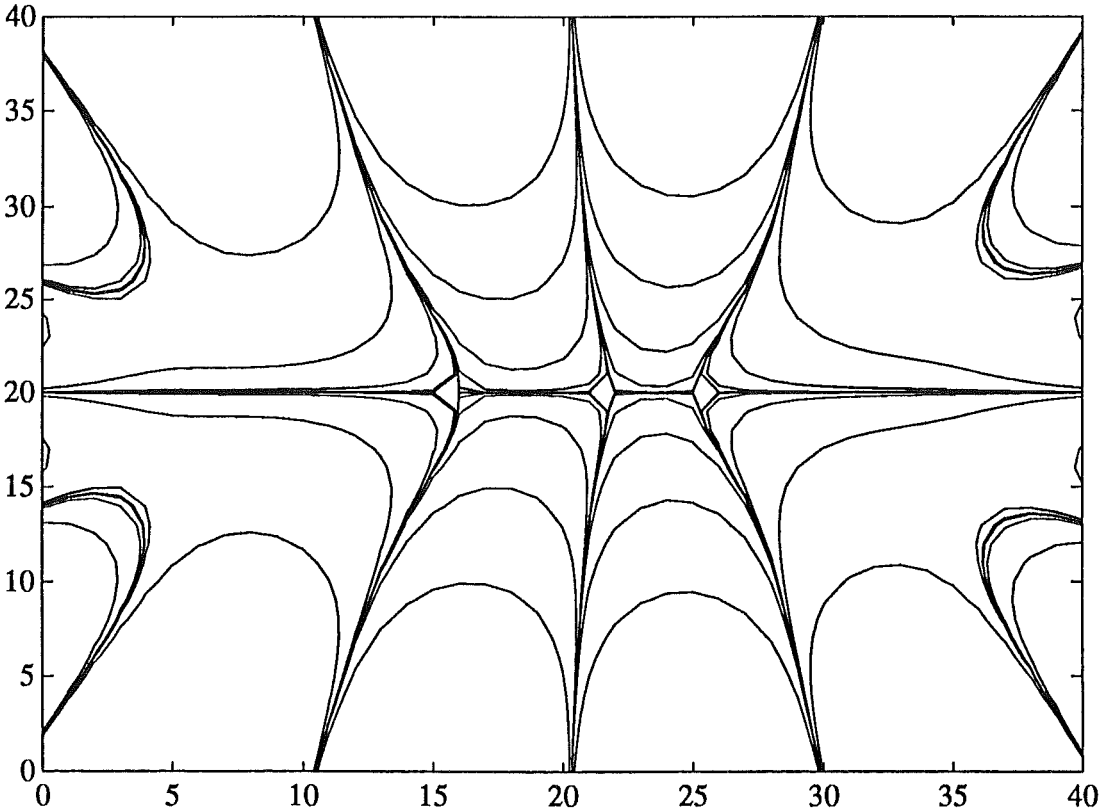


Figure 8 Contour map of Imag(e(z)).

2.2 Example 2 : $\log(1+x)$.

The $(4,4,4)$ approximation to $\log(1+x)$. Note that (see Chapter 4, 3.2.1)

(i)

$$\begin{aligned}&\left(6x^4-360x^3+180x^2+1080x+540\right)f(x)^2\\&+\left(-75x^4+1620x^3+5310x^2+3540x\right)f(x)\\&+260x^4-4080x^3-4080x^2=O\left(x^{14}\right)\end{aligned}$$

(ii) $y(x)=\dfrac{-a_1(x)+x\sqrt{d(x)}}{2a_2(x)}$ where

$$\begin{aligned}d(x)=&-615x^6+229320x^5-4136580x^4\\&+12612600x^3+59667300x^2+64033200x\\&+21344400.\end{aligned}$$

The roots of $d(x)$ are:

$$\begin{aligned}x&=354.0459\\x&=10.8301\pm 0.06444\mathbf{i}\\x&=-0.9155\pm 0.0005\mathbf{i}\\x&=-0.9972\end{aligned}$$

In Chapter 4 (3.2.1) it was shown that by defining a cut $\{x+\mathbf{i}y\in\mathbf{C}:x=-0.9155,|y|\leq 0.0005\}$ a good approximation to $\log(1+z)$ was obtained. The conjugate roots $z=10.8301\pm 0.06444\mathbf{i}$ can be treated in the same way. A cut has been defined as $\{x+\mathbf{i}y\in\mathbf{C}:x=10.8301,|y|\leq 0.06444\mathbf{i}\}$ and then the error function $e(z)$ graphed on the region $\{x+\mathbf{i}y\in\mathbf{C}:|x|,|y|\leq 20\}$ with a mesh spacing of 1.

Figures 9 and 10 are graphs of real$(e(z))$ and imag$(e(z))$ truncated at ± 1. Figures 11 and 12 are contour maps of real$(e(z))$ and imag$(e(z))$ with contours drawn at $\{\pm 1,\pm 10^{-1},\ldots,\pm 10^{-5}\}$.

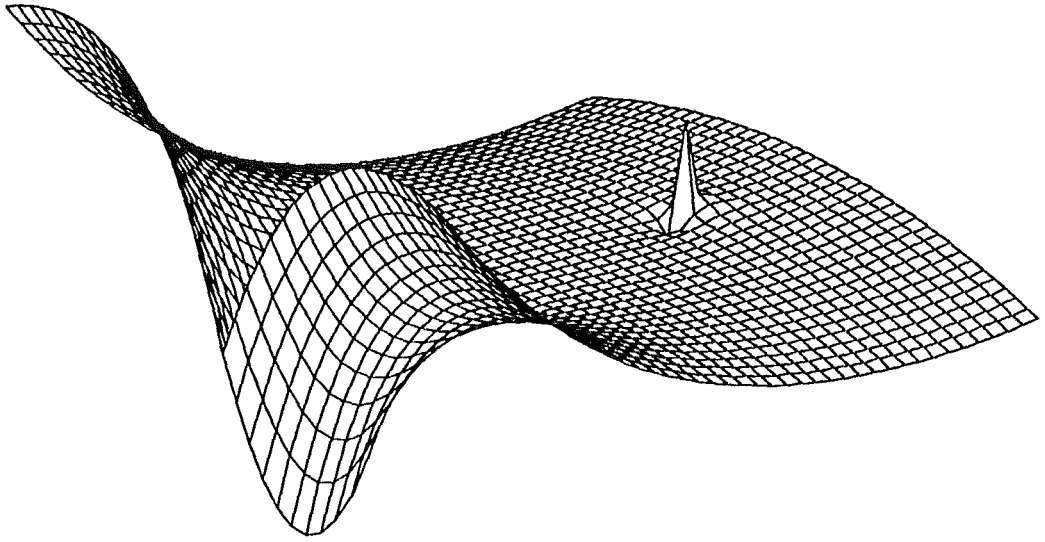


Figure 9 Real(e(z)). Truncation ± 1

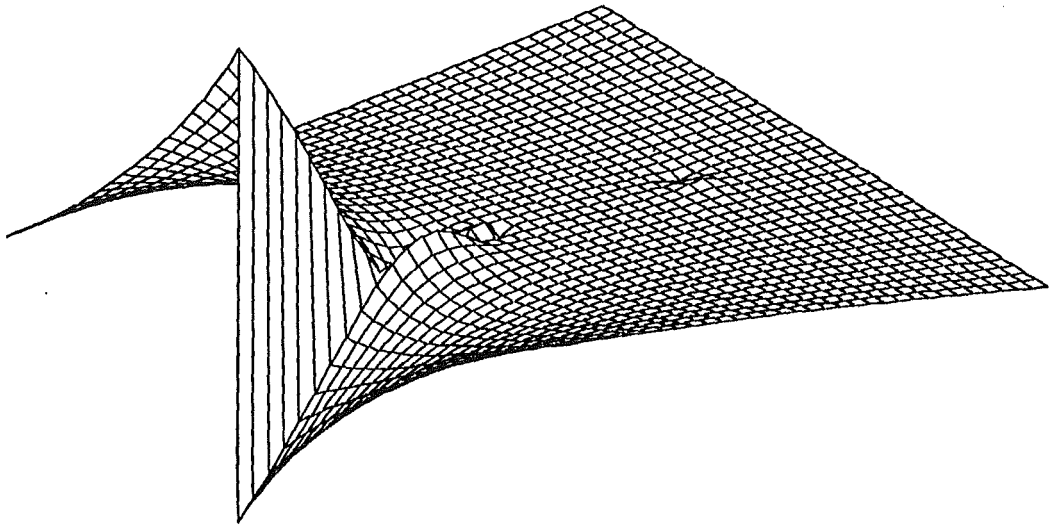


Figure 10 Imag(e(z)). Truncation ± 1

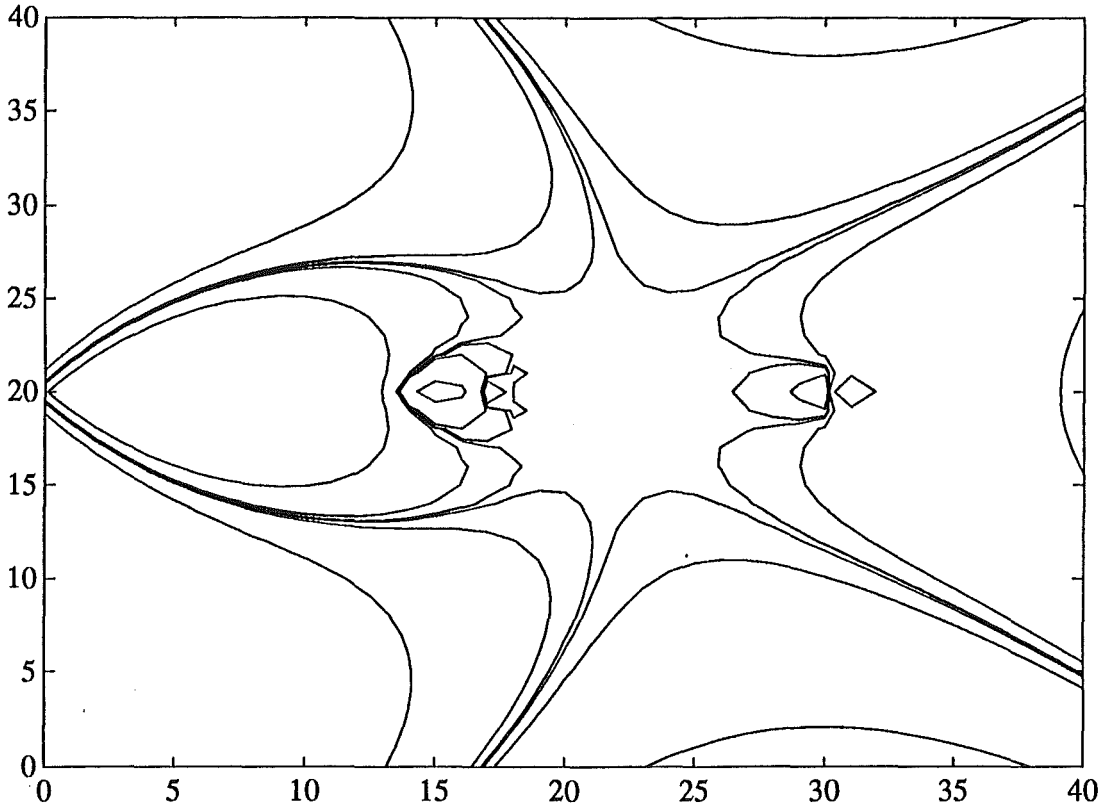


Figure 11 Contour map of Real(e(z)).

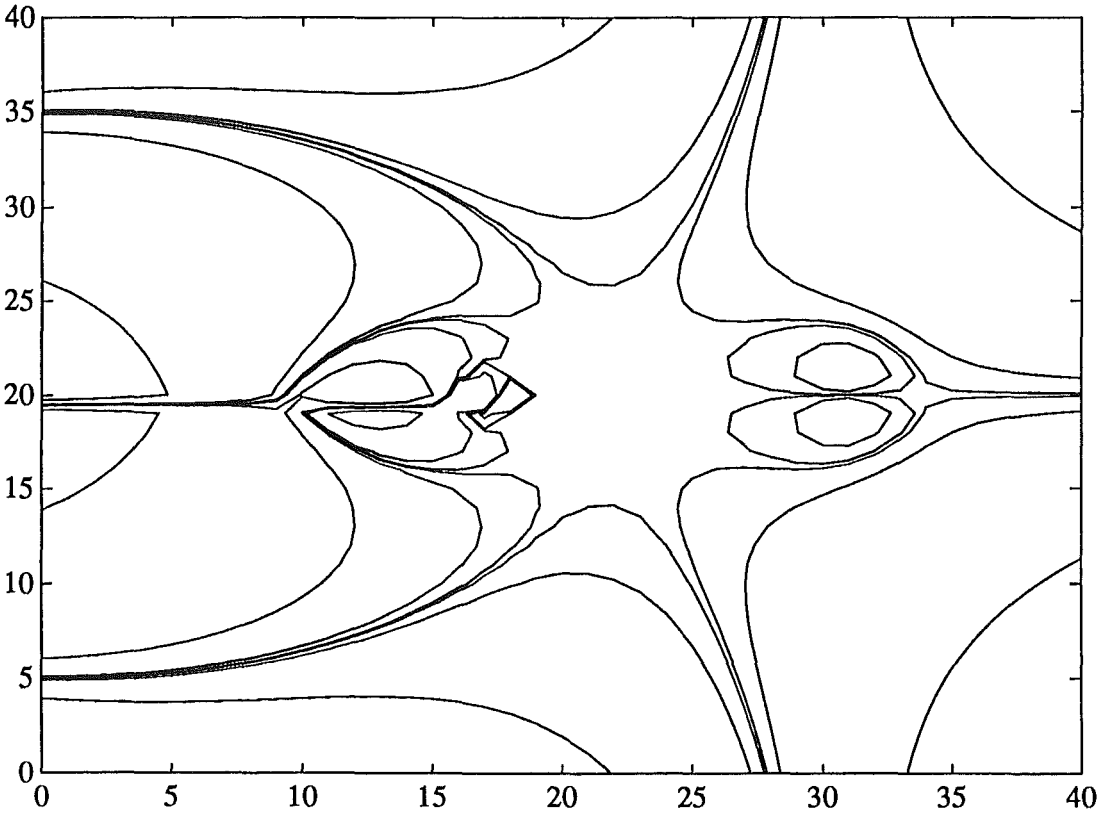


Figure 12 Contour map of Imag(e(z)).

The $(6,6)$ Padé approximation to $\log(1+x)$. Note that

$$p(x) = \frac{49x^6 + 1218x^5 + 7980x^4 + 20720x^3 + 23100x + 9240x}{10x^6 + 420x^5 + 4200x^4 + 16800x^3 + 31500x^2 + 27720x + 9240}$$

and that

$$y(x) = f(x) + O\left(x^{13}\right)$$
$$p(x) = f(x) + O\left(x^{13}\right).$$

Figures 13 and 14 are graphs of real$(e(z))$ and imag$(e(z))$ truncated at ± 1. Figures 15 and 16 are the usual contour maps of real$(e(z))$ and imag$(e(z))$.

Clearly $p(x)$ is inferior to $y(x)$ as an approximation to $\log(1+x)$ in this region. As a further illustration a graph showing $p(x), y(x)$, $\log(1+x)$along the positive real axis from 0 to 350 is given in Figure 17. Here $y(x)$ is represented by a solid line, $\log(1+x)$ by "– –" and $p(x)$ by "· · ·". $y(x)$ is reasonably accurate and certainly superior to $p(x)$ out to this point. As one approaches the branch point of $y(x)$ at $x = 354.0459$ the performance of $y(x)$ deteriorates. Beyond the branch point, however, real$(y(x))$ is still a good approximation.
e.g.

$$f(1000) = 6.91$$
$$\text{real}(y(1000)) = 6.50$$
$$p(1000) = 4.82$$

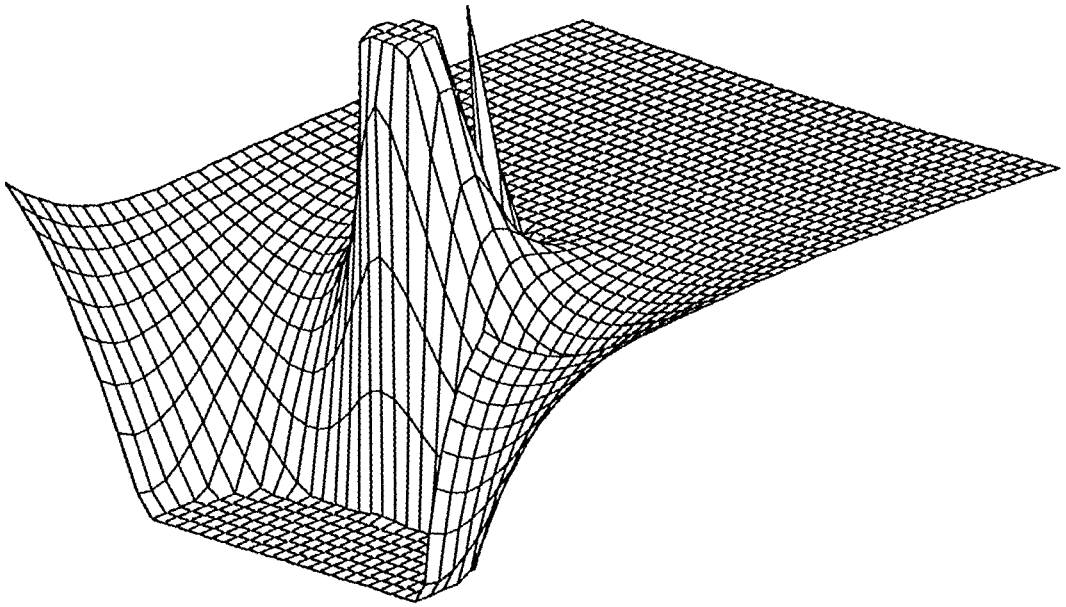


Figure 13 Real(e(z)). Truncation ± 1

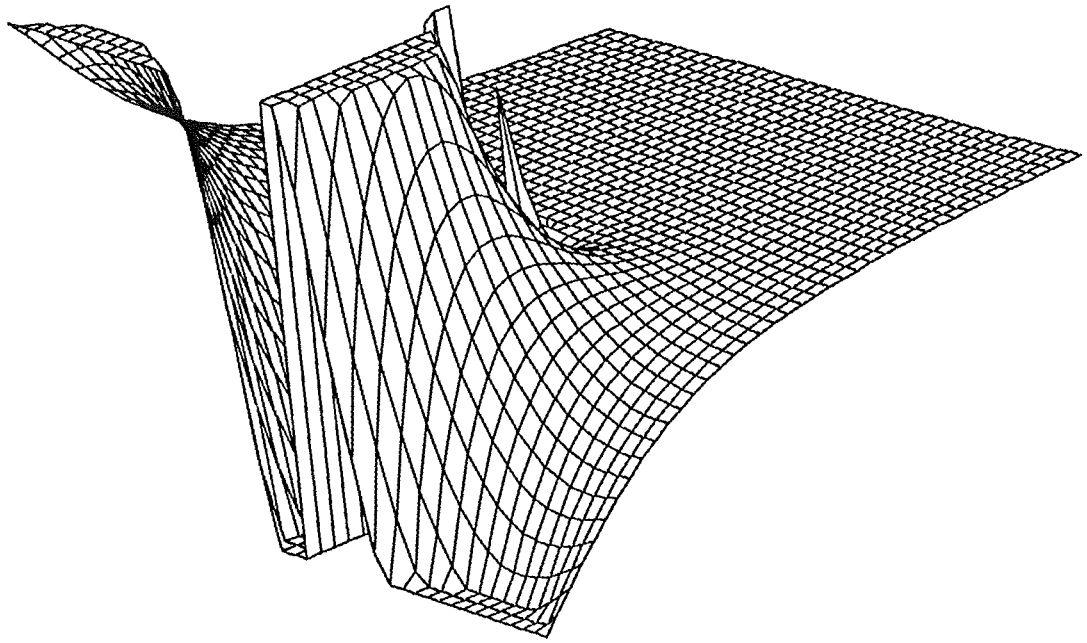


Figure 14 Imag(e(z)). Truncation ± 1

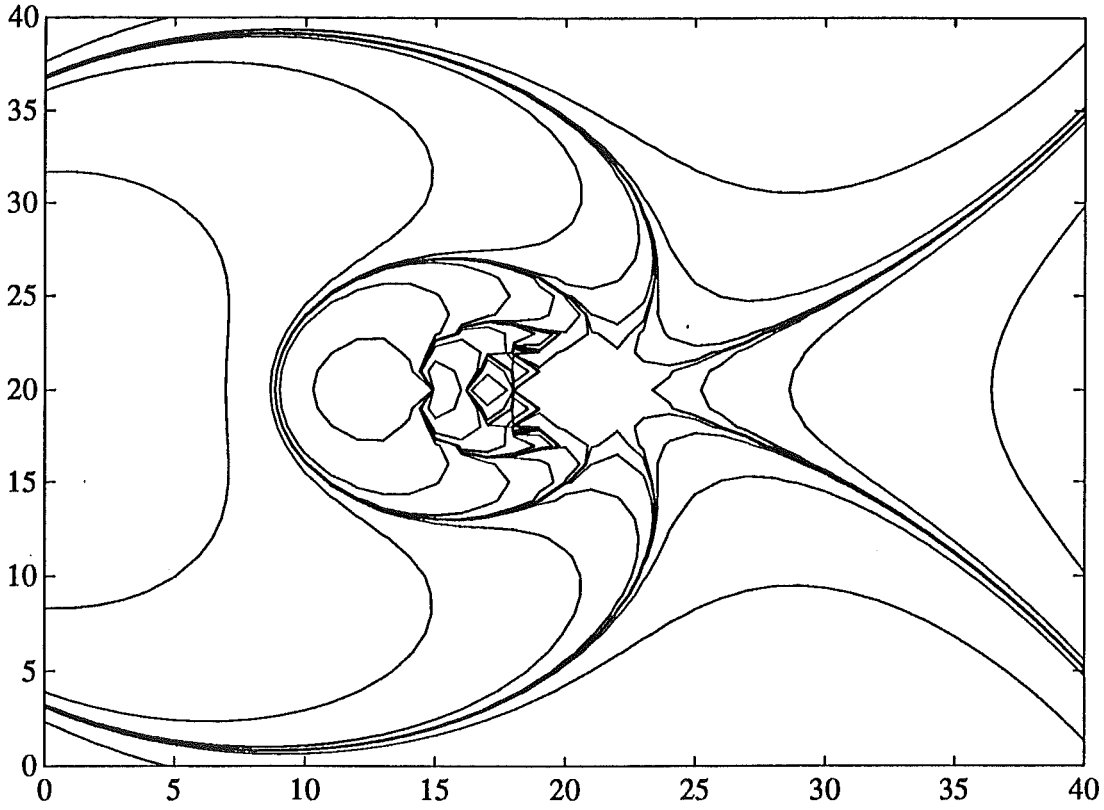


Figure 15 Contour map of Real(e(z)).

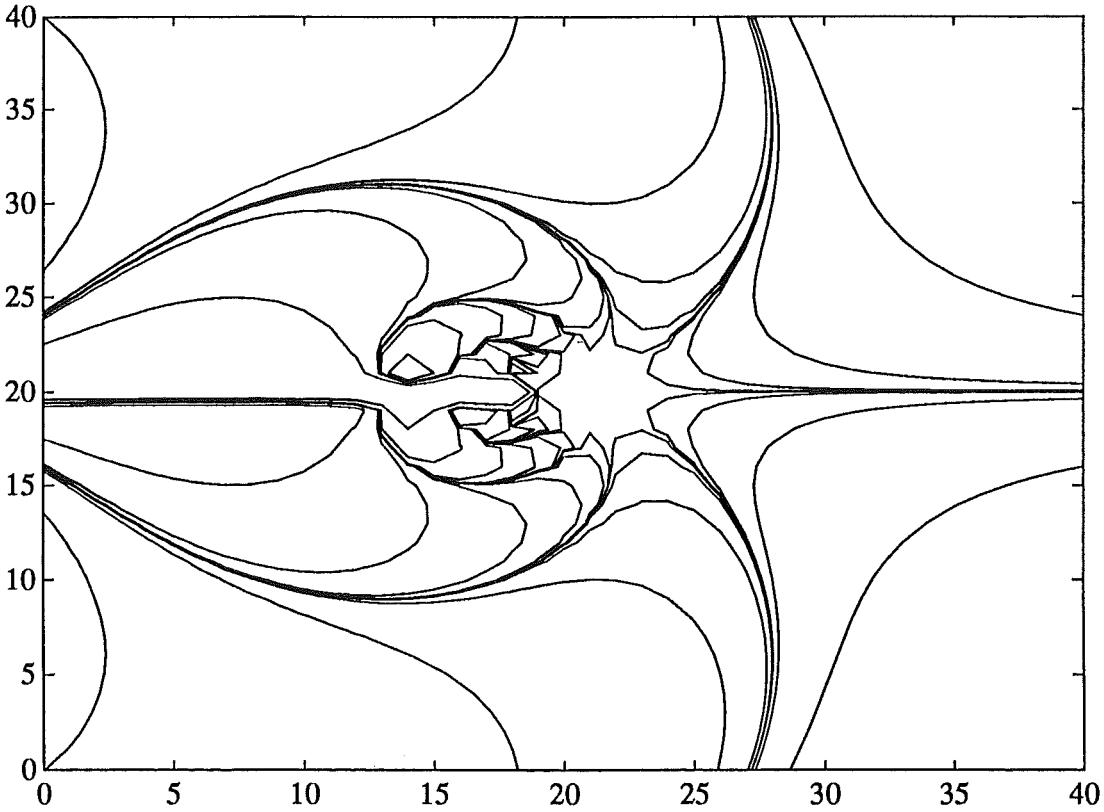


Figure 16 Contour map of Imag(e(z)).

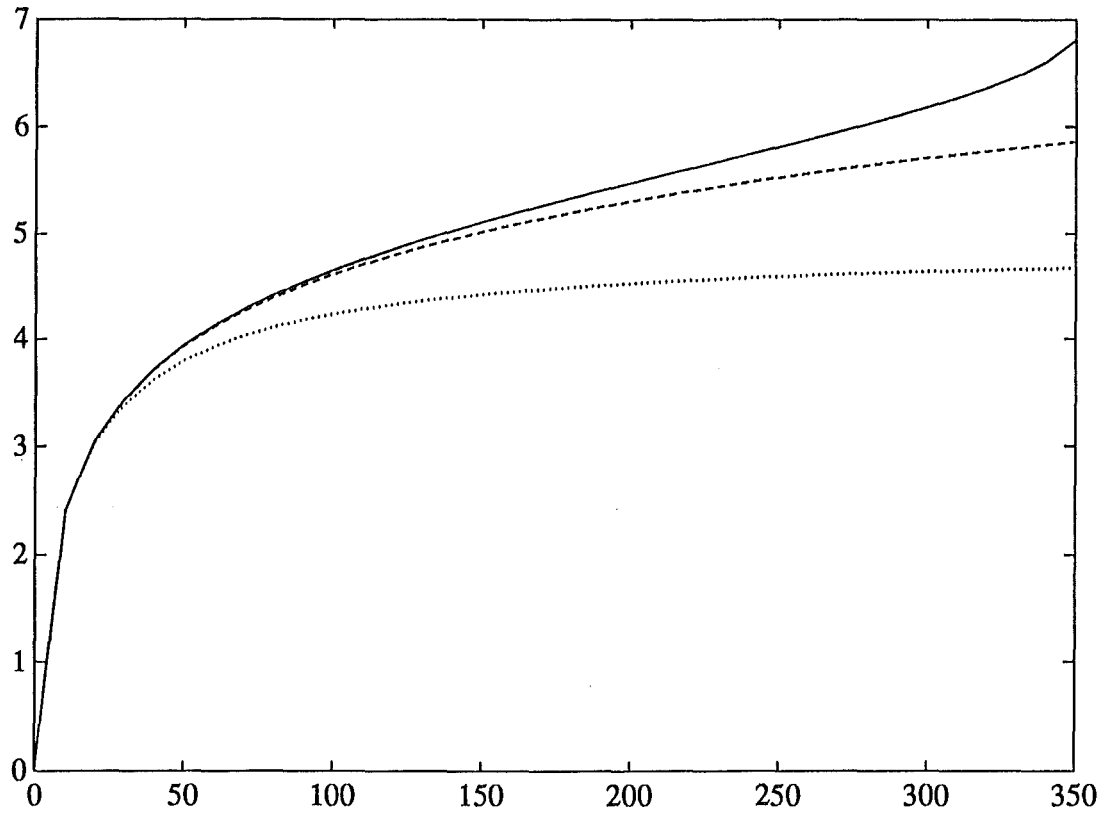


Figure 17

THE LIBRARY
UNIVERSITY OF CANTERBURY
CHRISTCHURCH, N.Z.

2.3 Example 3 : $\sqrt[3]{1+x}$.

The $(4,4,4)$ approximation to $\sqrt[3]{1+x}$. Note that

(i)

$$\begin{aligned}&\left(x^4-360x^3+1917x^2+2916x+729\right)f(x)^2\\+&\left(-14x^4+945x^3-513x^2-2916x-1458\right)f(x)\\+&\left(91x^4-1638x^3-2457x^2+729\right)\qquad=O\left(x^{14}\right)\end{aligned}$$

(ii)

$$y(x)=\frac{-a_1(x)+x\sqrt{d(x)}}{2a_2(x)}$$

where

$$\begin{aligned}d(x)=&-168x^6+111132x^5-2139291x^4\\&+7072758x^3+32470389x^2\\&+34720812x+11573604.\end{aligned}$$

The roots of $d(x)$ are:

$x=641.7609$

$x=11.2874\pm0.0393\mathbf{i}$

$x=-0.9186\pm0.0003\mathbf{i}$

$x=-0.9984$

Treating the roots of $d(x)$ as in the previous example and examining the region $\{x+\mathbf{i}y\in\mathbf{C}:|x|,|y|\leq 20\}$ with a mesh spacing of 1 it can be seen that $y(x)$ is a good approximation to $\sqrt[3]{1+x}$ with a similar branch point structure.

Figures 18 and 19 are graphs of real$(e(z))$ and imag$(e(z))$ truncated at ±1. Figures 20 and 21 are the usual contour maps of real$(e(z))$ and imag$(e(z))$.

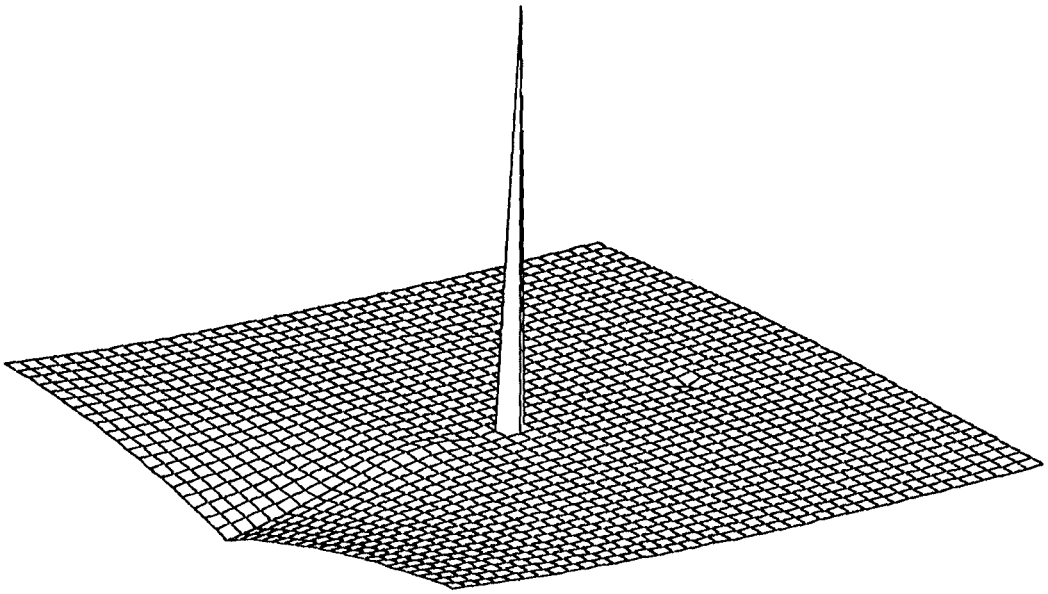


Figure 18 Real(e(z)). Truncation ± 1

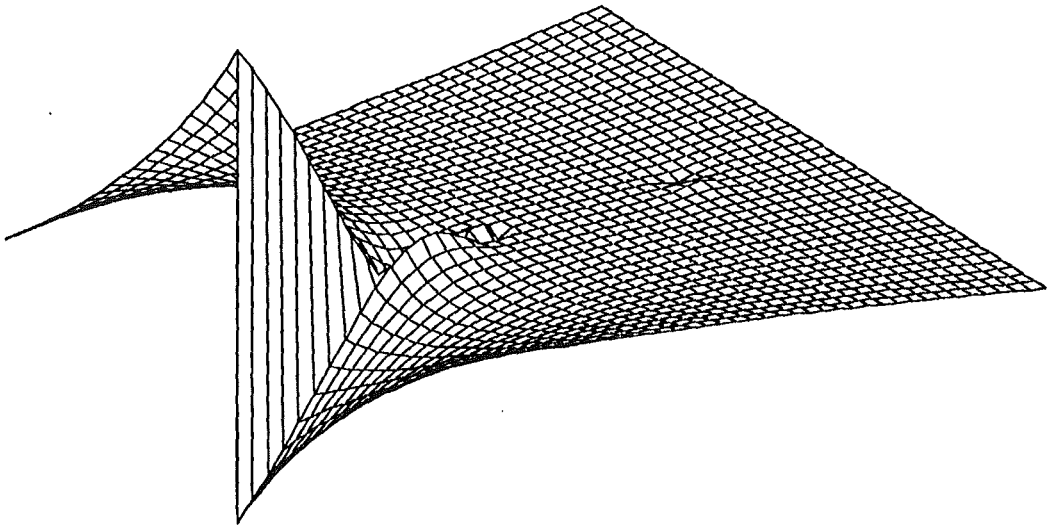


Figure 19 Imag(e(z)). Truncation ± 1

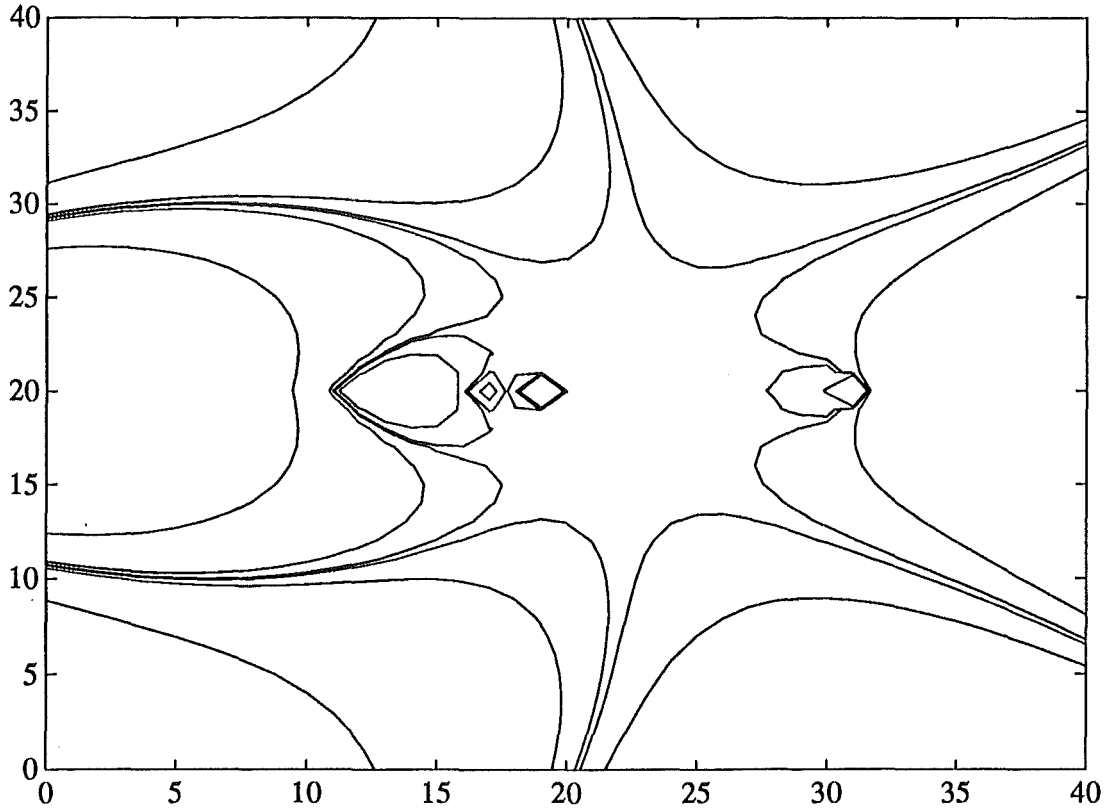


Figure 20 Contour map of Real(e(z)).

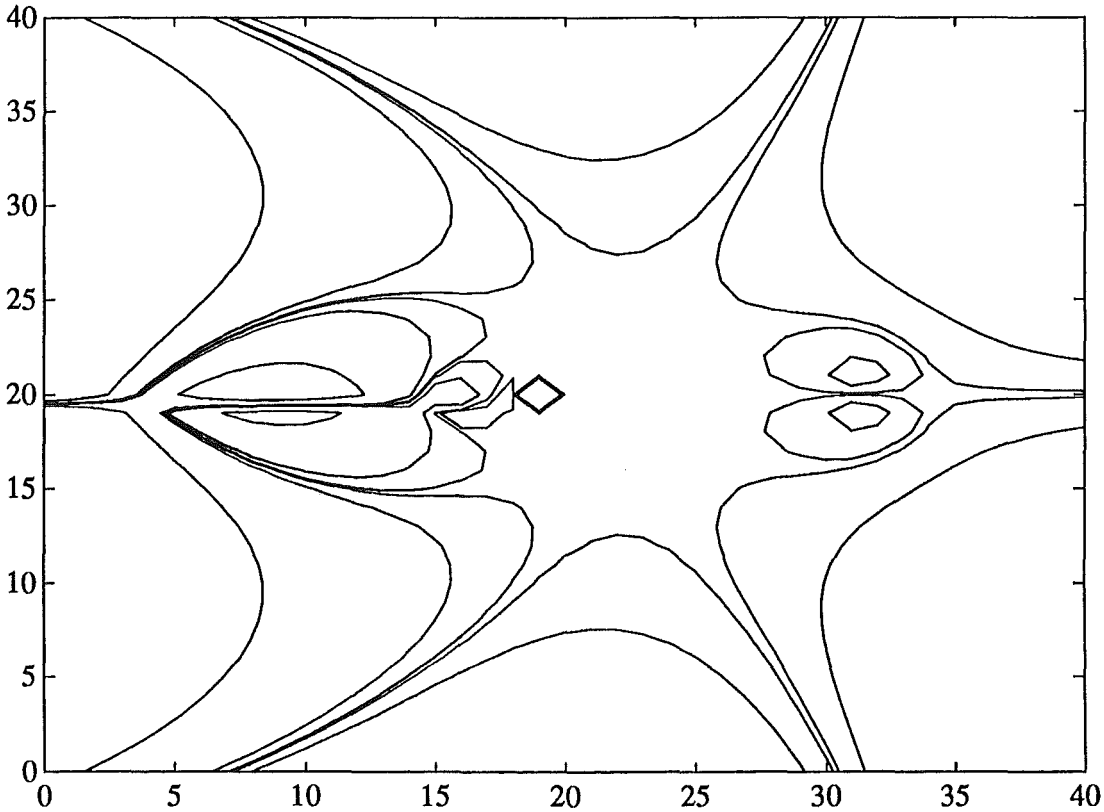


Figure 21 Contour map of Imag(e(z)).

The $(6,6)$ Padé approximation to $\sqrt[3]{1+x}$. Note that

$$p(x) = \frac{988x^6 + 31122x^5 + 266760x^4 + 960336x^3 + 1662120x^2 + 1371249x + 433026}{187x^6 + 11781x^5 + 141372x^4 + 636174x^3 + 1301265x^2 + 1226907x + 433026}$$

and that

$$y(x) = f(x) + O\left(x^{13}\right)$$
$$p(x) = f(x) + O\left(x^{13}\right) .$$

Figures 22 and 23 are graphs of real$(e(z))$ and imag$(e(z))$ truncated at ± 1. Figures 24 and 25 are the usual contour maps of real$(e(z))$ and imag$(e(z))$.

One can draw here the same conclusions as in the previous example. Again, as a further illustration a graph showing $p(x), y(x), \sqrt[3]{1+x}$ along the positive real axis from 0 to 600 is given in Figure 26. Here $y(x)$ is represented by a solid line, $\sqrt[3]{1+x}$ by "– –" and $p(x)$ by "...". Again beyond the branch point real$(y(x))$ is still a good approximation.
e.g.

$$f(999) = 10$$
$$\text{real}(y(999)) = 10.17$$
$$p(999) = 5.12$$

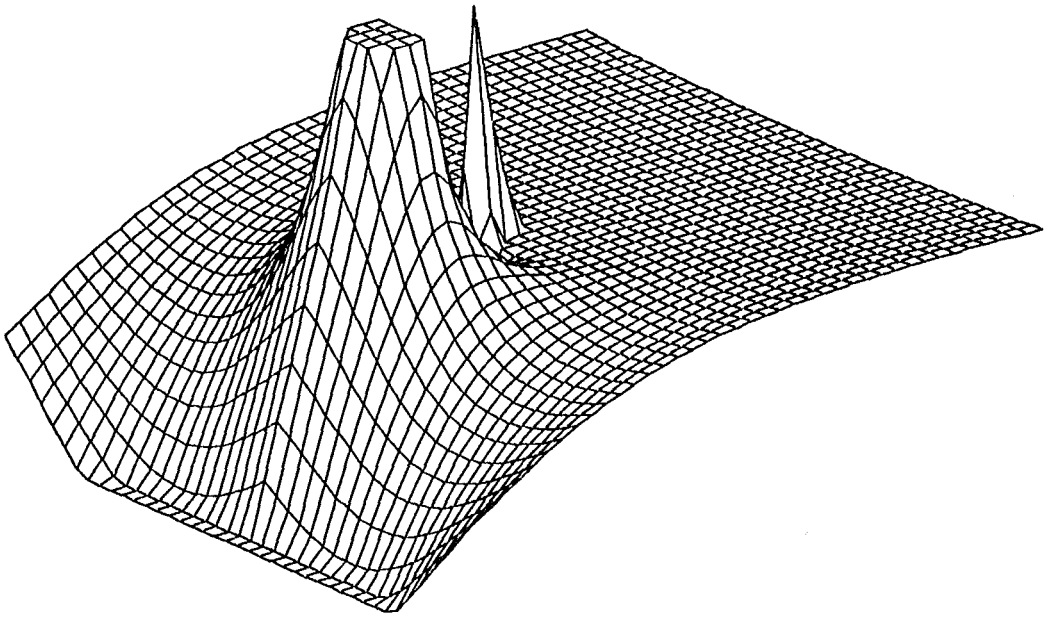


Figure 22 Real(e(z)). Truncation ± 1

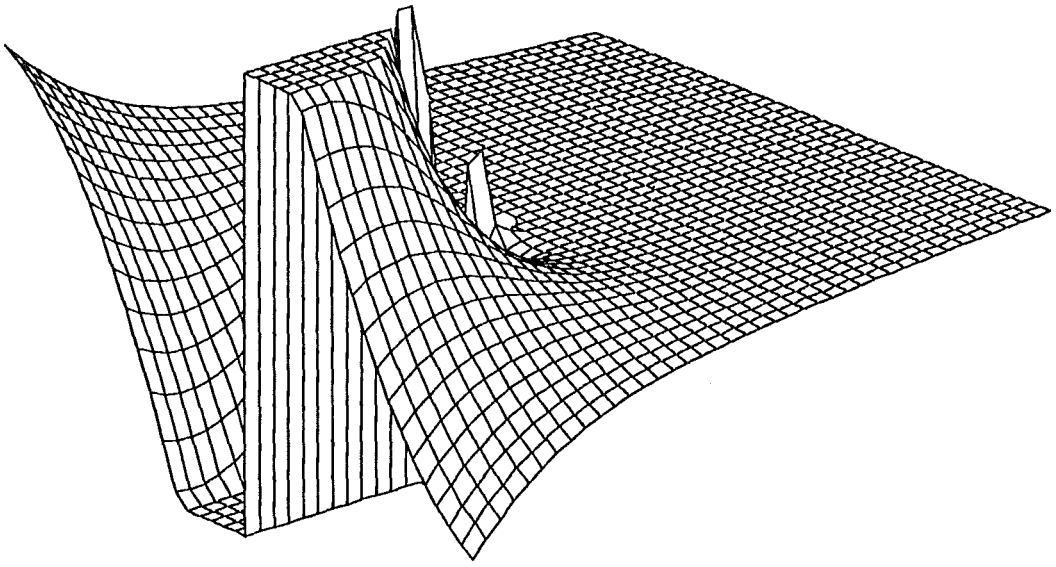


Figure 23 Imag(e(z)). Truncation ± 1

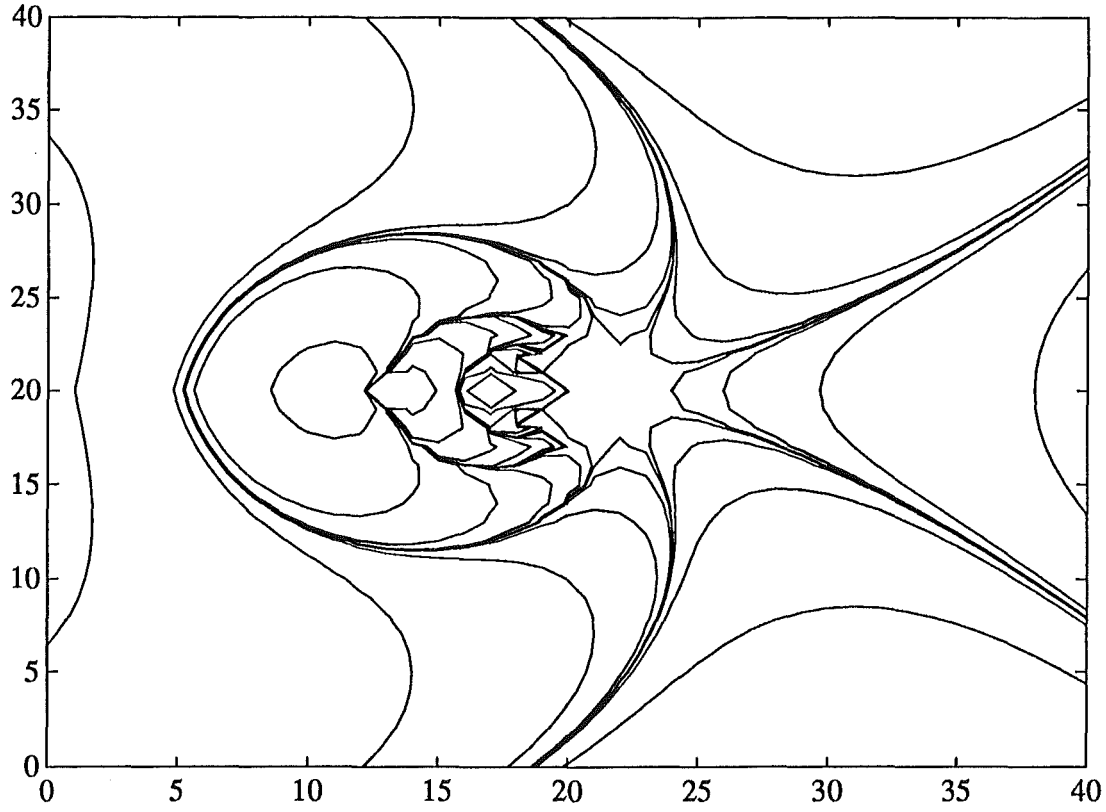


Figure 24 Contour map of Real(e(z)).

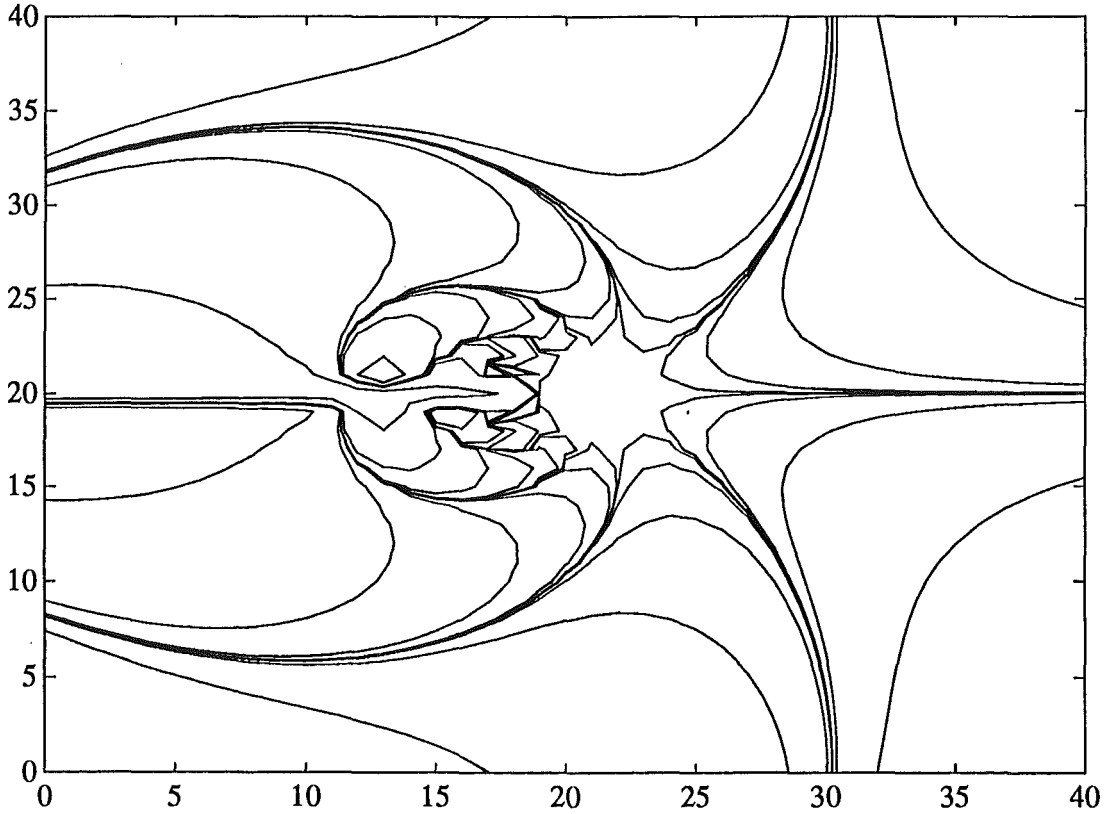


Figure 25 Contour map of Imag(e(z)).

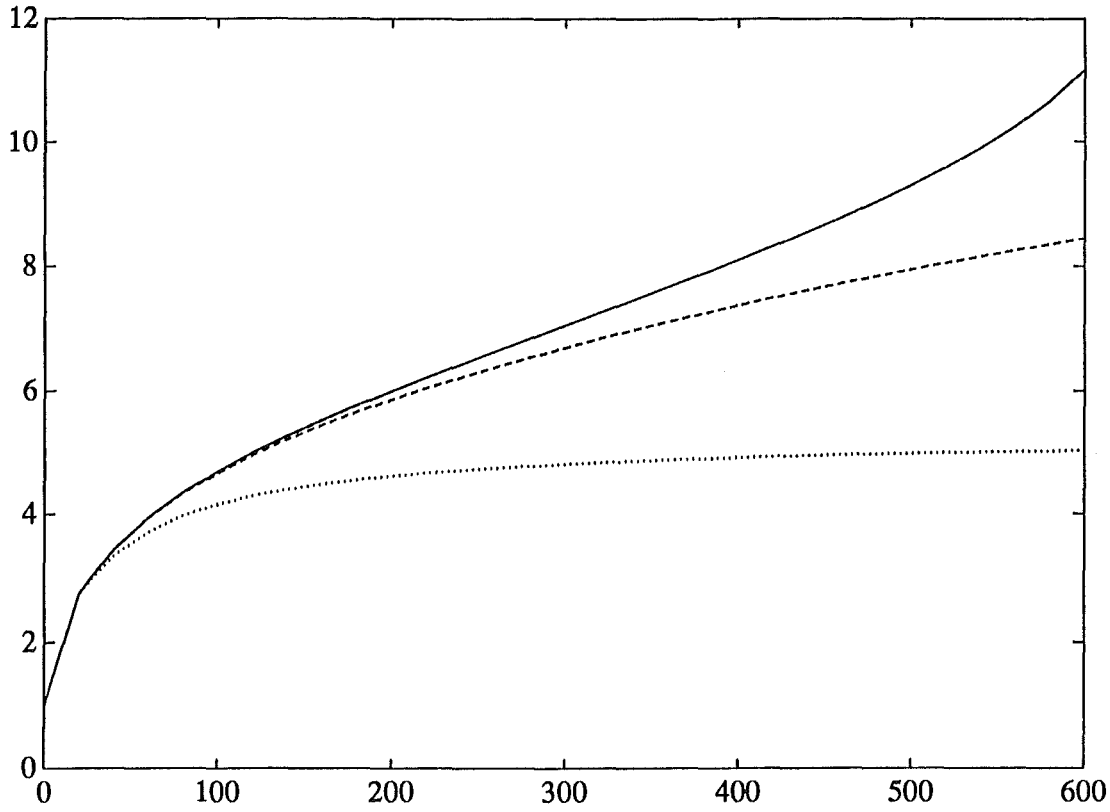


Figure 26

3. Conclusion

In each of these examples the area over which the quadratic approximation performs well seems, at first, to be limited by points at which $D(x) = 0$. If, however, these points are zeroes of "small" separation this has been shown not to be the case. The quadratic approximation may be extended beyond these points giving an approximation which is significantly better than the Padé approximation over a wide area.

CHAPTER 6

SEQUENCES AND STRUCTURE

1. Introduction

Up to this point the existence and behaviour of individual quadratic Hermite-Padé forms and approximations have been considered. The questions of the structure of the quadratic table and the behaviour of sequences of approximations are still to be addressed. The following topics are considered in this chapter:

(i) In §2 it is shown that almost any sequence $\{(A_2^i, A_1^i, A_0^i) : A_j^i, i \in \mathbf{N}\}$ with $\lim_{i\to\infty} \sum_j A_j^i = \infty$ gives sequences of quadratic approximations $y_i(x) = f(x) + O(x^{n_i})$ with $\lim_{i\to\infty} n_i = \infty$. .

(ii) In §3 it is shown that there is a best choice from a multidimensional space of (A_2, A_1, A_0) forms.

(iii) In §4 a simple proof of the block structure of the Padé table is sketched. Generalisations of this idea are shown to give an indication of the structure of quadratic Hermite-Padé tables. Several examples are given.

2. Sequences of quadratic Hermite-Padé approximations

Let $R = \max\{r \in \mathbf{N} : f(x) = O(x^r)\}$ and let $\{(A_2^i, A_1^i, A_0^i) : A_j^i, i \in \mathbf{N}\}$ be a sequence of number triples satisfying:

(i) $A_2^i + A_1^i + 1 \geq R$

(ii) $(A_0^i + A_1^i)/2 \geq R$

(iii) $\lim_{i\to\infty} \sum_j A_j^i = \infty$.

(i) and (ii) are minor conditions, avoiding undue degeneracy in any corresponding sequence of Hermite-Padé forms (see the following lemma) while (iii) ensures that such a sequence has increasing order. Let $\{(a_2^i(x), a_1^i(x), a_0^i(x)), i \in \mathbf{N}\}$ be the polynomial coefficients of a sequence of Hermite-Padé forms corresponding to
$\{(A_2^i, A_1^i, A_0^i) : i \in \mathbf{N}\}$.

Lemma. $\forall i \in \mathbf{N}$ at least two members of $\{a_2^i(x), a_1^i(x), a_0^i(x)\}$ are not identically zero.

Proof. If $a_2^i(x) \equiv 0 \equiv a_1^i(x)$ then

$$a_0^i(x) = O\left(x^{A_0^i+2}\right)$$

which is impossible.

If $a_2^i(x) \equiv 0 \equiv a_0^i(x)$ then

$$a_1^i(x) f(x) = O\left(x^{A_2^i + A_1^i + A_0^i + 2}\right)$$
$$\Rightarrow f(x) = O\left(x^{A_2^i + A_0^i + 2}\right)$$

which contradicts (i) above.

If $a_1^i(x) \equiv 0 \equiv a_0^i(x)$ then

$$a_2^i(x) f(x)^2 = O\left(x^{A_2^i + A_1^i + A_0^i + 2}\right)$$
$$\Rightarrow f(x)^2 = O\left(x^{A_1^i + A_0^i + 2}\right)$$
$$\Rightarrow f(x) = O\left(x^{(A_1^i + A_0^i + 2)/2}\right)$$

which contradicts (ii) above. □

The following theorem shows that such an "increasing" sequence of Hermite-Padé forms yields a sequence of approximations with increasing order of accuracy.

Theorem 1. Let $\{[(A_2^i, A_1^i, A_0^i)] : i \in \mathbf{N}\}$ be a sequence of Hermite-Padé forms with the above properties. Then this sequence gives a sequence of quadratic approximations $\{y_i(x) : i \in \mathbf{N}\}$ to $f(x)$ such that:

(i) $y_i(x) = f(x) + O(x^{n_i})$

(ii) $\lim_{i \to \infty} n_i = \infty$.

Proof. Let $i \in \mathbf{N}, r_i = \max\{r \in \mathbf{N} : x^r | a_j^i(x), \forall j \in \{0,1,2\}\}$, $a_j^i(x) = x^r g_j^i(x)$. Then

$$\sum_j g_j^i(x) f(x)^j = O\left(x^{N_i + 2 - r_i}\right) \qquad (\text{where } N_i = \sum_j A_j^i).$$

Using the results of Chapter 3, there exists $y_i(x)$, a quadratic (possibly rational or polynomial) approximation such that

$$y_i(x) = f(x) + O\left(x^{(N_i + 2 - r_i)/2}\right)$$

Since at least two of the $a_j^i(x)$ are not identically zero and $x^{r_i} | a_j^i(x) \quad \forall j \in \{0,1,2\}$ then $N_i \geq 2r_i$.

Consider the sequence $\{N_i - r_i : i \in \mathbf{N}\}$. Noting that $N_i - r_i \geq N_i - \frac{N_i}{2}$ and that $\lim_{i \to \infty} N_i = \infty$ it follows that $\lim_{i \to \infty} (N_i - r_i) = \infty$. Letting $n_i = (N_i - r_i + 2)/2$, the result is proven.

□

Although this result falls far short of proving any sort of convergence it at least gives hope that such a general result may exist. Convergence results so far published are rather restricted and certainly avoid any kind of degeneracy (see Baker and Lubinsky [2]).

3. An optimal choice from the space of (A_2, A_1, A_0) forms

In general, the space of (A_2, A_1, A_0) forms for $f(x)$ may be multi-dimensional. This corresponds to the matrix appearing in equation (3) of Chapter 1 having rank < N+2. If the rank is N+3–k then the space of forms has dimension k. If k>1 there is a choice of which solution to take. In the context of a sequence of Hermite-Padé forms Theorem 1 indicates that it is not too important which solution is chosen. If, however, a basis for the solution space is known it is clearly desirable to have some method of choosing a form of maximal order from this space. That is, we seek a one dimensional subspace whose elements satisfy $\sum a_i(x) f(x)^i = O\left(x^M\right)$ where M is maximal over the space of (A_2, A_1, A_0) forms. A procedure for finding such a subspace is now given.

Let S_{k_1} be the space of all (A_2, A_1, A_0) forms for $f(x)$ and let $\left\{m^{i,w} : i \epsilon \{1, ..., k_w\}\right\}$ be a basis for the k_w dimensional space S_{k_w} (to be defined below) of (A_2, A_1, A_0) forms.
Let $\left\{a_j^{i,w}(x) : j \epsilon \{0,1,2\}\right\}$ be the polynomial coefficients of $m^{i,w}$ and
let $\left\{e_j^{i,w} : j \epsilon \{0,1,2,...\}\right\}$ be the coefficients of the power series $\sum_{j=0}^{\infty} e_j^{i,w} x^j = \sum_{j=0}^{2} a_j^{i,w}(x) f(x)^j$

Step 1 :Set $w = 1$

Step 2 :If $k_w = 1$:

then the one dimensional subspace has been found so terminate.

Else if $\sum e_j^{i,w} x^j \equiv 0$ for some i

then $f(x)$ has been found exactly so terminate.

Else set $S_{k_{w+1}} = \left\{ \sum_{i=1}^{k_w} c_i m^{i,w} : \sum_{i=1}^{k_w} c_i e_{N+1+w}^{i,w} = 0 \right\}$

Step 3 :Set $w = w + 1$.

Go to Step 2.

Note that:

(i) $S_{k_{w+1}}$ is clearly a linear subspace of S_{k_w} (so $k_{w+1} \leq k_w$) and it is non-trivial since $\sum_{i=1}^{k_w} c_i e_{N+1+w}^{i,w} = 0$ is a linear homogeneous equation in the k_w (≥ 2) unknowns c_i.

(ii) This process must terminate since if $k_w > 1 \quad \forall w \epsilon \mathbf{N}$ then $\exists p \epsilon \mathbf{N}$ such that $\sum_{p=0}^{\infty} e_j^{1,p} x^j \equiv 0$.

(iii) Since this procedure has raised the order of the form (A_2, A_1, A_0) by w after w iterations, it follows that $m^{i,w}$ is also a form of types

$$\left\{(A_2 + r, A_1 + s, A_0 + t) \ : \ r, s, t \, \epsilon \, \mathbf{Z}^+, \ r + s + t \leq w - 1\right\}.$$

One common example of when this may occur is when $f(x)$ is an even function. If $f(x)$ is even then so is $f(x)^2$, and an examination of the matrix appearing equation (3) of Chapter 1 in this case reveals that if the A_k are all even then the matrix has rank of at most $N + 1$. Hence, in this case the solution space for an even function has a dimension of at least two. This can be observed in the tables of quadratic forms for $\cos(x)$ given in Chapter 1 and later in this chapter.

Example.

A basis for the $(0, 2, 0)$ forms for $\cos(x)$ is given by (see Chapter 1 Table 2):

$m^{1,1} = (x^2 + 2)\, y - 2, \quad e_4^{1,1} = -5/12$

$m^{2,1} = y^2 - 2y + 1 \qquad e_4^{2,1} = 1/4$

So $m^{1,2} = 3m^{1,1} + 5m^{2,1} = 5y^2 + (3x^2 - 4)y - 1$ (with $5f(x)^2 + (3x^2 - 4)f(x) - 1 = O(x^5)$).

Note that this means that $m^{1,2}$ is a form of types $(1, 2, 0), (0, 3, 0)$ and $(0, 2, 1)$. This can be seen in Table 2 of Chapter 1.

In fact $5f(x)^2 + (3x^2 - 4)f(x) - 1 = O(x^6)$ so it is also a form of other types also.

4. Structure and degeneracy in the table of quadratic Hermite-Padé forms.

4.1 The Padé table

To provide some indication of possible generalisations to the quadratic case we first give some results concerning the Padé table. A D-table showing the degenerate Padé forms is introduced instead of the more traditional C-table (see [1]). This table gives more information about the structure and degeneracy of the Padé forms and leads directly to the Padé table of approximants in much the same way as the C-table. These results (and parts of the proofs thereof) are adapted from Baker [1].

Let $\{m_i : i \in I\}$ be the linear space of (A_1, A_0) Padé forms (for $f(x)$). Let $m_i = a_1^i(x)y + a_0^i(x)$. Note that if $a_1^i(0) = 0$ then $a_0^i(0) = 0$ (since $a_1^i(x)f(x) + a_0^i(x) = O(x^{A_1+A_0+1})$). Let $r_i = \max\{r \in \mathbf{N} : x^r | a_j^i(x) \quad \forall j \in \{0,1\}\}$ and let $a_j^i(x) = x^{r_i} b_j^i(x)$. Then $y(x) = \dfrac{-b_0^i(x)}{b_1^i(x)} = f(x) + O\left(x^{A_1+A_0-r_i+1}\right)$ and $b_1^i(0) \neq 0$.

Theorem 2. (Uniqueness) (see Baker [1] Theorem 1.1).
If $y_i(x)$ and $y_j(x)$ are the Padé approximants corresponding to the (A_1, A_0) Padé forms m_i, m_j, then

$$y_i(x) = y_j(x) \qquad \forall i, j \in I.$$

Proof.

$$\frac{b_0^i(x)}{b_1^i(x)} - \frac{b_0^j(x)}{b_1^j(x)} = O\left(x^{A_1+A_0+1-max\{r_i,r_j\}}\right)$$

$$\Rightarrow b_0^i(x)\, b_1^j(x) - b_0^j(x)\, b_1^i(x) = O\left(x^{A_1+A_0+1-\max\{r_i,r_j\}}\right) \tag{1}$$

But the left hand side of (1) is a polynomial of degree at most $A_0 - r_i + A_1 - r_j < A_1 + A_0 + 1 - \max\{r_i, r_j\}$ thus is identically zero. Hence

$$\frac{b_0^i(x)}{b_1^i(x)} \equiv \frac{b_0^j(x)}{b_1^j(x)}$$

and it follows that all (A_1, A_0) Padé forms are just polynomial multiples of some basic form.

□

From Theorem 2, one can choose a unique (up to a constant factor) representative of the (A_1, A_0) forms, i.e. a form of minimal degree from those of maximal order.

Note that :

(i) $p(A_1, A_0)$ has no common polynomial factor except possibly $x^j, j \in \mathbf{N}$.

(ii) If $D(A_1, A_0) < \infty$ then one of the coefficients of $p(A_1, A_0)$ must have full degree.

(iii) In the case $D(A_1, A_0) < \infty$ the conditions (i) and (ii) above completely characterise $p(A_1, A_0)$.

Let $R(m_i)(x) = \sum_{j=0}^{1} a_j^i(x) f(x)^j$, for $m_i \in S$, the linear space of (A_1, A_0) forms.

Definition 1 :The unique representative, $p(A_1, A_0)$ of the (A_1, A_0) forms is defined by

$$p(A_1, A_0) = \Big\{ m_i : \sum_j \deg\left(a_j^i(x)\right) = \min\Big\{ \sum_j \deg(a_j^k(x)) : O(R(m_k)(x)) = \max\Big\{ O(R(m_r)(x)) : r \in I \Big\} \Big\} \Big\}$$

Definition 2 :The degeneracy, $D(A_1, A_0)$ of the Padé form $p(A_1, A_0)$ is defined by

$$D(A_1, A_2) = \mathrm{Ord}(R(p(A_1, A_2))(x)) - (A_1 + A_0 + 1)$$

where $\mathrm{Ord}(R(x)) = N$ if $O(R(x)) = O(x^N), \neq O(x^{N+1})$ as $x \to \infty$.

The degeneracy, D, is the amount of extra matching obtained from $p(A_1, A_0)$. Trefethen [3] introduced a similar concept for the Padé approximation rather than the Padé form, when studying a related problem. However this definition and presentation makes these concepts clearer. The two dimensional table of D values, the D-table, will be shown to give the block structure of the Padé table in a somewhat easier fashion than the C-table (see Baker [1], p13).

Theorem 3. Let $D(A_1, A_0) < \infty$. Then

$$p(A_1, A_0) = x^r p(A_1 - r, A_0 - r), r \in \mathbf{N} \Leftrightarrow D(A_1 - r, A_0 - r) = D(A_1, A_0) + r .$$

Proof. Note that $D(A_1 - r, A_0 - r) = D(A_1, A_0) + r$

$$\Leftrightarrow \mathrm{Ord}(R(p(A_1, A_0))(x)) = \mathrm{Ord}(R(p(A_1 - r, A_0 - r))(x)) + r .$$

If $p(A_1, A_0) = x^r p(A_1 - r, A_0 - r)$ then clearly $D(A_1 - r, A_0 - r) = D(A_1, A_0) + r$.
If $D(A_1 - r, A_0 - r) = D(A_1, A_0) + r$ then

$$\mathrm{Ord}(R(p(A_1, A_0))(x)) = \mathrm{Ord}(R(p(A_1 - r, A_0 - r))(x)) + r .$$

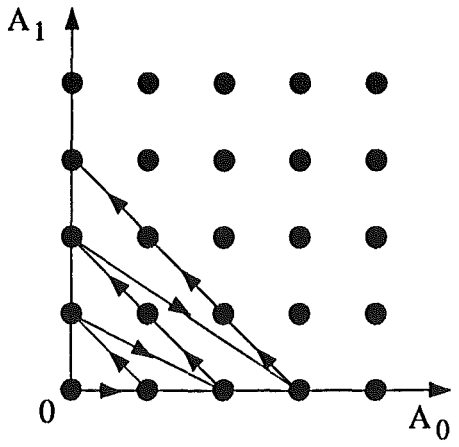
Certainly $\mathrm{Ord}(R(x^r p(A_1 - r, A_0 - r))(x)) = \mathrm{Ord}(R(p(A_1, A_0))(x))$. Since
$x^r p(A_1 - r, A_0 - r)$ has no common factor except $x^s, s \in \mathbf{N}$ then
$p(A_1, A_0) = x^r p(A_1 - r, A_0 - r)$. □

It is obvious that $p(A_1, A_0)$ yields an approximation

$$y(x) = f(x) + O\left(x^{A_1+A_0-r+D(A_1,A_0)+1}\right) \tag{2}$$

where x^r is the maximal factor of $p(A_1, A_0)$. So the only "problems" in the Padé table occur when $r > D(A_1, A_0)$. In short, using Theorem 3, a form gives an approximation with less than expected order of accuracy if and only if some appropriate earlier form has $D > 0$.

Now set out the Padé table and search through forms in the order shown (this is similar to Baker's proof of the structure of the Padé table ([1],p17)).



Note in passing that the forms on the borders of the table can easily be shown to have no factor of x. Suppose $p(A_1, A_0)$ is the first form found with $D(A_1, A_0) = t > 0$. The table is "normal" (see Baker [1], p24) up to this point. $p(A_1, A_0)$ generates a block structure (we take for example $t = 3$). In the table that follows $[p(A_1, A_0), D]$ will denote $p(A_1, A_0)$ with $D(A_1, A_0) = D$. This table is rotated 90 degrees from the traditional forms in Baker [1]. Note how the equal D values lie rows and columns bordering the initial element as in the diagram below.

0	0	0	0
1	1	1	0
2	2	1	0
3	2	1	0

$[p(A_1+4,A_0),D(A_1+4,A_0)]$	$[p(A_1+4,A_0+1),D(A_1+4,A_0+1)]$	$[p(A_1+4,A_0+2),D(A_1+4,A_0+2)]$	$[p(A_1+4,A_0+3),D(A_1+4,A_0+3)]$	$[p(A_1+4,A_0+4),D(A_1+4,A_0+4)]$
$[p(A_1,A_0),0]$	$[xp(A_1,A_0),0]$	$[x^2p(A_1,A_0),0]$	$[x^3p(A_1,A_0),0]$	$[p(A_1+3,A_0+4),D(A_1+3,A_0+4)]$
$[p(A_1,A_0),1]$	$[xp(A_1,A_0),1]$	$[x^2p(A_1,A_0),1]$	$[x^2p(A_1,A_0),0]$	$[p(A_1+2,A_0+4),D(A_1+2,A_0+4)]$
$[p(A_1,A_0),2]$	$[xp(A_1,A_0),2]$	$[xp(A_1,A_0),1]$	$[xp(A_1,A_0),0]$	$[p(A_1+1,A_0+4),D(A_1+1,A_0+4)]$
$[p(A_1,A_0),3]$	$[p(A_1,A_0),2]$	$[p(A,A_0),1]$	$[p(A_1,A_0),0]$	$[p(A_1,A_0+4,D(A_1,A_0+4)\ ]$

Table 1

If $a_1(x)$ and $a_0(x)$ are the polynomial coefficients of $p(A_1, A_0)$ then since $D(A_1-1, A_0) = D(A_1, A_0-1) = 0$ it follows that $\deg(a_j(x)) = A_j$. Also, $p(A_1, A_0)$ does not have a factor of x since, by assumption, $D(A_1-1, A_0-1) = 0$. The entries in the table are now justified.

(i) The entries $\{(A_1+i, A_0+j) : i,j \in \{0,1,2,3\}\}$:
For a given (i,j) the entry is easily seen to be a (A_1+i, A_0+j) form with no common factor except possibly $x^j, j \in \mathbf{N}$. Also one of the coefficients of the entry has maximum degree so this entry equals $p(A_1+i, A_0+j)$ (by note (iii) after Theorem 2).

(ii) The top and right borders of entries do not have a factor of x:
Take for example (A_1+4, A_0) and suppose otherwise. Then $p(A_1+4, A_0) = xp(A_1+3, A_0-1)$ and $D(A_1+3, A_0-1) > 0$. Since $p(A_1+3, A_0-1)$ has no common factor except possibly $x^t, t \in \mathbf{Z}^+$ then $p(A_1+3, A_0) = x^j p(A_1+3, A_0-1), j \in \mathbf{Z}^+$. But $p(A_1+3, A_0) = p(A_1, A_0)$ and $p(A_1, A_0)$ has no factor of x so $p(A_1+3, A_0-1) = p(A_1, A_0)$. Hence $\deg(a_0(x)) < A_0$ a contradiction.

The final structure is that of a 3×3 block of forms which have $x^r, r \in \mathbf{N}$ as a factor. The forms $\{(A_1+1+i, A_0+1+j) : i+j \leq D(A_1, A_0)-2, \quad i,j \in \mathbf{Z}^+\}$ still yield an approximation of full order whilst the others in the block do not. In Baker's terminology (Baker [1]) these latter approximations are said "not to exist" although clearly they do, but with a less than expected order of accuracy. A Padé table having such blocks is said to be "non-normal".

The concept of the D-table outlined above, generalises easily to give the standard theorem on the structure of the Padé table approximations. The version quoted below is a modification of that from Baker [1,Theorem 2.3].

Theorem 4. The Padé table can be completely dissected into $r \times r$ blocks with horizontal and vertical sides, $r \geq 1$. Let $[\lambda/\mu]$ denote the unique ($\lambda + \mu =$ minimum) member of a particular $r \times r$ block. Then:

(i) The $[\lambda/\mu]$ exists and the numerator and denominator are of full nominal degree.

(ii) $[\lambda+p/\mu+q] = [\lambda/\mu]$ for $p+q \leq r-1, p \geq 0, q \geq 0$.

(iii) $[\lambda+p/\mu+q] = [\lambda/\mu]$ for $p+q \geq r, r-1 \geq p \geq 1, r-1 \geq q \geq 1$, but as an approximation of less than full order, given by (2).

(iv) $C(\lambda+p/\mu+q) = 0$ for $1 \leq p \leq r-1, 1 \leq q \leq r-1$ and $C \neq 0$ otherwise ($C(\lambda+p/\mu+q)$ is $a_1(0)$ in our notation).

4.2 The quadratic Hermite-Padé table

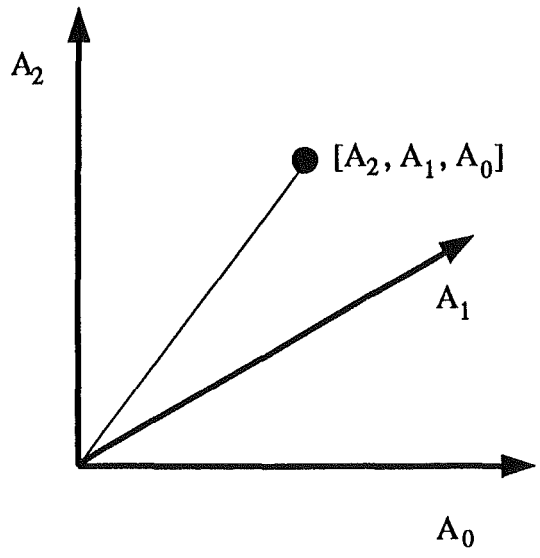
In the previous section it was shown that the Padé table has a simple structure and one would hope that similar theorems could be formulated in the quadratic case. The basic property which underlies the Padé structure is that of Theorem 2, i.e. that any two (A_1, A_0) Padé forms differ by at most a polynomial factor. This is *not*, unfortunately, true for quadratic forms as can be seen by the example in §3. Hence, although, as will be shown, a table of D values gives much valuable information, complications may arise and break down any systematic structure.

Basic degenerate structures

If a form has a degeneracy $D > 0$ then this propagates to the forms around it in a manner similar to that in the Padé table. Examples are given below, the various entries being deduced from simple order matching principles. It should be realised that, lacking any kind of uniqueness theorem, each entry is *one* (i, j, k) form only. There may be others, which may be of different degeneracy. Unless a type of uniqueness result can be found it is expected that the general case will involve overlapping of the following structures. This approach does, however, completely explain the structure of many of the previous examples.

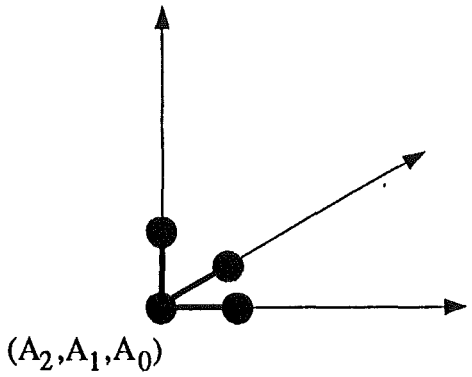
Examples

As usual, the following $3 - D$ representation is used:



Example 1

$D(A_2, A_1, A_0) = 1$

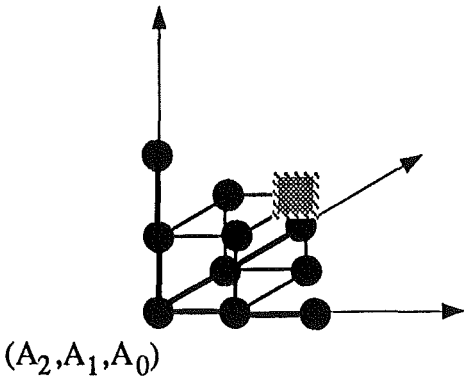


● denotes $[A_2,A_1,A_0]$

For instance, see the example at the end of §3 where $5y^2 + (3x^2 - 4)y - 1$ is a (0,2,0) form for $\cos x$ giving this structure, i.e. the (0,2,0) form is also a form of types (1,2,0), (0,3,0), (0,2,1).

Example 2

$D(A_2, A_1, A_0) = 2$



● denotes $[A_2,A_1,A_0]$

▒ denotes $x[A_2,A_1,A_0]$

The (0,0,0) form for $\cos(x)$, that is $y^2 - 2y + 2$, gives this structure (see Chapter 1 Table 2).

A further example is given in the tables following. These give bases for the relevant solution spaces for $\cos(x)$. Taking simple linear combinations of the basis elements, the (2,2,2) form, $(23x^2+465)y^2+(344x^2+960)y+578x^2-1425$, can be seen to generate this structure; yet overlapping it starting at (2,2,4) is another such structure generated by the form $(14x^2+417)y^2+(704x^2+11136)y-237x^4+5267x^2-11553$ (see the comments immediately preceding this section of examples).

Table 2: Hermite-Padé forms for cos(x)

$A_2 = 2$

A_1 / A_0	2	3	4
2	$a_0(x) = 11x^2 - 27$ $a_1(x) = 4x^2 + 24$ $a_2(x) = 3$	$a_0(x) = 578x^2 - 1425$ $a_1(x) = 344x^2 + 960$ $a_2(x) = 23x^2 + 465$	$a_0(x) = -723x^2 + 1783$ $a_1(x) = -23x^4 - 12x^2 - 2096$ $a_2(x) = 313$
	$a_0(x) = -49x^2 + 120$ $a_1(x) = -12x^2 - 120$ $a_2(x) = x^2$		$a_0(x) = 22483x^2 - 55440$ $a_1(x) = 465x^4 + 4924x^2 + 55440$ $a_2(x) = 313x^2$
3	$a_0(x) = 578x^2 - 1425$ $a_1(x) = 344x^2 + 960$ $a_2(x) = 23x^2 + 465$	$a_0(x) = 578x^2 - 1425$ $a_1(x) = 344x^2 + 960$ $a_2(x) = 23x^2 + 465$	$a_0(x) = -5828x^2 + 14379$ $a_1(x) = -474x^4 + 5152x^2 - 28128$ $a_2(x) = 361x^2 + 13749$
4	$a_0(x) = -23x^4 + 477x^2 - 1037$ $a_1(x) = 48x^2 + 1024$ $a_2(x) = 13$	$a_0(x) = -237x^4 + 5267x^2 - 11553$ $a_1(x) = 704x^2 + 11136$ $a_2(x) = 14x^2 + 417$	$a_0(x) = 361x^4 - 8367x^2 + 18447$ $a_1(x) = -28x^4 - 768x^2 - 18624$ $a_2(x) = 177$
	$a_0(x) = 465x^4 - 9137x^2 + 20160$ $a_1(x) = -766x^2 - 20160$ $a_2(x) = 13x^2$		$a_0(x) = -4583x^4 + 105269x^2 - 231840$ $a_1(x) = 278x^4 + 10592x^2 + 231840$ $a_2(x) = 59x^2$

$A_2 = 3$

A_0 \ A_1	2	3	4
2	$a_0(x) = 578x^2 - 1425$ $a_1(x) = 344x^2 + 960$ $a_2(x) = 23x^2 + 465$	$a_0(x) = 578x^2 - 1425$ $a_1(x) = 344x^2 + 960$ $a_2(x) = 23x^2 + 465$	$a_0(x) = -5828x^2 + 14379$ $a_1(x) = -474x^4 + 5152x^2 - 28128$ $a_2(x) = 361x^2 + 13749$
3	$a_0(x) = 578x^2 - 1425$ $a_1(x) = 344x^2 + 960$ $a_2(x) = 23x^2 + 465$	$a_0(x) = 578x^3 - 1425x$ $a_1(x) = 344x^3 + 960x$ $a_2(x) = 23x^3 + 465x$	$a_0(x) = -5828x^2 + 14379$ $a_1(x) = -474x^4 + 5152x^2 - 28128$ $a_2(x) = 361x^2 + 13749$
4	$a_0(x) = -237x^4 + 5267x^2 - 11553$ $a_1(x) = 704x^2 + 11136$ $a_2(x) = 14x^2 + 417$	$a_0(x) = -237x^4 + 5267x^2 - 11553$ $a_1(x) = 704x^2 + 11136$ $a_2(x) = 14x^2 + 417$	$a_0(x) = 17911x^4 - 427753x^2 + 946395$ $a_1(x) = -2416x^4 - 26944x^2 - 984960$ $a_2(x) = 782x^2 + 38565$

$A_2 = 4$

A_0 \ A_1	2	3	4
2	$a_0(x) = 578x^2 - 1425$ $a_1(x) = 344x^2 + 960$ $a_2(x) = 23x^2 + 465$	$a_0(x) = 74095x^2 - 182802$ $a_1(x) = 66112x^2 + 69504$ $a_2(x) = 237x^4 + 7843x^2 + 113298$	$a_0(x) = -5828x^2 + 14379$ $a_1(x) = -474x^4 + 5152x^2 - 28128$ $a_2(x) = 361x^2 + 13749$
	$a_0(x) = -519x^2 + 1279$ $a_1(x) = -216x^2 - 1088$ $a_2(x) = x^4 - 191$		$a_0(x) = 305727x^2 - 754287$ $a_1(x) = 15686x^4 - 69792x^2 + 1036704$ $a_2(x) = 361x^4 - 282417$
3	$a_0(x) = 74095x^2 - 182802$ $a_1(x) = 66112x^2 + 69504$ $a_2(x) = 237x^4 + 7843x^2 + 113298$	$a_0(x) = 74095x^2 - 182802$ $a_1(x) = 66112x^2 + 69504$ $a_2(x) = 237x^4 + 7843x^2 + 113298$	$a_0(x) = -3571907x^2 + 8813250$ $a_1(x) = -745936x^4 + 13104064x^2 - 39012480$ $a_2(x) = 17911x^4 + 1160833x^2 + 30199230$
4	$a_0(x) = -237x^4 + 5267x^2 - 11553$ $a_1(x) = 704x^2 + 11136$ $a_2(x) = 14x^2 + 417$	$a_0(x) = -46621x^4 + 1083296x^2 - 2389095$ $a_1(x) = 180608x^2 + 2234880$ $a_2(x) = 151x^4 + 7751x^2 + 154215$	$a_0(x) = 17911x^4 - 427753x^2 + 946395$ $a_1(x) = -2416x^4 - 26944x^2 - 984960$ $a_2(x) = 782x^2 + 38565$
	$a_0(x) = 7843x^4 - 169923x^2 + 371523$ $a_1(x) = -19392x^2 - 364416$ $a_2(x) = 14x^4 - 7107$		$a_0(x) = -50471x^4 + 1198575x^2 - 2650095$ $a_1(x) = 5392x^4 + 100800x^2 + 2701440$ $a_2(x) = 34x^4 - 51345$

These results may be summarised in the following theorem.

Theorem 5. If a (A_2, A_1, A_0) form has degeneracy $D(A_2, A_1, A_0) = t \in \mathbf{N}$ then :

$$x^r\,[A_2, A_1, A_0]\,,\quad r\epsilon\left\{0, 1, ..., max\left\{s:\ s \leq t/2,\ s\epsilon\mathbf{Z}^+\right\}\right\}$$

is a form of type

$$\left\{(A_2 + r + i_r, A_1 + r + j_r, A_0 + r + k_r)\ :\ i_r + j_r + k_r \leq t - 2r,\ \ i_r, j_r, k_r\epsilon\mathbf{Z}^+\right\}$$

with

$$D\,(A_2 + r + i_r, A_1 + r + j_r, A_0 + r + k_r) = t - 2r - (i_r + j_r + k_r)\ .$$

Proof. Since

$$R([A_2, A_1, A_0])\,(x) = O\left(x^{A_2+A_1+A_0+2+t}\right)$$

then

$$R(x^r\,[A_2, A_1, A_0])\,(x) = O\left(x^{A_2+A_1+A_0+2+t+r}\right)$$

Hence $x^r[A_2, A_1, A_0]$ will be a quadratic form of type $(A_2 + r, A_1 + r, A_0 + r)$ provided that $r \leq t/2$. The degeneracy of this form is $t - 2r$. Further, it follows, as in §3, that this form is also a form of the type

$$\left\{(A_2 + r + i_r, A_1 + r + j_r, A_0 + r + k_r)\ :\ i_r + j_r + k_r \leq t - 2r,\ \ i_r, j_r, k_r\epsilon\mathbf{Z}^+\right\}$$

with

$$D\,(A_2 + r + i_r, A_1 + r + j_r, A_0 + r + k_r) = t - 2r - (i_r + j_r + k_r)\ .$$

□

Note also that there is a direct connection between k, the dimension of the solution space for (A_2, A_1, A_0) forms and the degeneracy, D, of the optimal (see §3) (A_2, A_1, A_0) form, namely $D \geq k - 1$. The example $f(x) = \cos(x), (A_2, A_1, A_0) = (0, 0, 0)$ (see Chapter 1, Table 2) shows that equality need not hold.

5. Conclusion

In this chapter several miscellaneous topics have been explored :

(i) In §2 it was shown that a sequence of Hermite-Padé forms of increasing degree gives a sequence of approximations of increasing order. This result depends only on several weak assumptions about the sequence and $O(f(x))$.

(ii) In §3 it was shown that there is a 1 dimensional subspace of each space of (A_2, A_1, A_0) forms whose elements are of maximal order. This, of course, applies to all Hermite-Padé systems. A form of maximal order does not necessarily yield an approximation of the same order, as can be seen by the example of the (0,0,0) approximation to $\cos(x)$.

(iii) In §4 the question of the structure of the table of quadratic forms was examined. Firstly we introduced the concept of the D-table of Padé forms and deduced the well-known theorem on the block structure of the Padé table. This idea was applied to the D-table for quadratic forms, and whilst not completely solving the problem, because of the complication of overlapping structures, it was shown to indicate the sorts of structures which occur.

6. References

1. G.A. Baker (1975): Essentials of Padé Approximants. New York: Academic Press.
2. G.A. Baker, D.S. Lubinsky (1987): *Convergence theorems for rows of differential and algebraic Hermite-Padé approximations.* J. Comput. Appl. Math., **18** : 29–52.
3. L.N. Trefethen (1984): *Square blocks and equioscillation in the Padé, Walsh and CF tables* in Rational Approximation and Interpolation, P.R. Graves-Morris et al. (eds). New York: Springer Verlag.

CHAPTER 7

SUMMARY

In this thesis we have attempted to solve, as completely as possible, the questions associated with the quadratic Hermite-Padé approximation. At this, the final stage, it is appropriate to summarise the extent to which this has been achieved.

The starting point was the definition of the general Hermite-Padé form of type $(A_n, \ldots, A_0)$ for the system $g_0(f(x)), \ldots, g_n(f(x))$ as, briefly, a non-trivial set of polynomials $\{a_i(x) : \deg(a_i(x)) \leq A_i, \ i \in \{0, \ldots, n\}\}$ such that

$$\sum_{i=0}^{i} a_i(x)\, g_i(f(x)) = O\left(x^{N+n}\right) \quad \text{where} \quad N = \sum_{i=0}^{n} A_i\,. \tag{1}$$

An approximation $y(x)$ is derived from this form by setting

$$\sum_{i=0}^{n} a_i(x)\, g_i(y(x)) \equiv 0\,. \tag{2}$$

The first question which arises is concerned with the practical calculation of such forms. This was a source of unexpected difficulty. Attempts to solve, numerically, the system of equations represented by (1) were unsuccessful since this system was found to be, in most cases, almost singular.

A partial solution to this problem was found by using the symbolic arithmetic of MACSYMA. This produces Hermite-Padé forms for systems with rational coefficients and provided a large number of concrete examples.

In Chapter 2 another approach yielded a recursive algorithm. Although this algorithm was given explicitly for the diagonal quadratic case it is clear that the same approach will work in any more general case which is sufficiently normal. This algorithm is based on simple order-matching principles rather than complicated identities and produces sequences such as $\{[(1,1,1)], \ldots, [(n,n,n)]\}$ in $O(n^2)$ operations which compares extremely favourably with the $O(n^4)$ operations required to produce such a sequence using Gaussian elimination on each system. Unfortunately, the algorithm is numerically unstable, another manifestation of the original difficulty.

The second question arising immediately from the definition is this. When does (2) define a $y(x)$ which is a useful approximation to $f(x)$? In the case of algebraic Hermite-Padé approximations this was shown to depend on the conditions:

(i) $a_n(0) \neq 0$

(ii) $\frac{\partial}{\partial y}\left(\sum_{i=0}^{n} a_i(x)\, y^i\right)\Big|_{x=0} \neq 0\,.$

In Chapter 3 it was shown that with the quadratic approximation these conditions are unnecessary.

In this case an explicit expression for $y(x)$ can be given as

$$y(x) = \frac{-a_1(x) \pm \sqrt{D(x)}}{2a_2(x)} \quad \text{where} \quad D(x) = a_1(x)^2 - 4a_2(x)\,a_0(x)\ .$$

Condition (i) above is readily shown to be unimportant, however if $D(0) = 0$, $y(x)$ does not at first appear to be analytic at the origin. In fact, if $D(x) = x^{2s}g(x)$ $\quad, g(0) \neq 0$, then rewriting $y(x)$ as

$$y(x) = \frac{-a_1(x) \pm x^s\sqrt{g(x)}}{a_2(x)}$$

solves this problem. It was shown that $D(x)$ is always of the form which allows this to be done. The conclusion of Chapter 3 is that if $D(0) = 0$ then $D(x) = x^{2s}g(x)$, as above, and $y(x)$ is an approximation of slightly degraded order. In detail, given $a_2(x)f(x)^2 + a_1(x)f(x) + a_0(x) = O\left(x^{N+2}\right)$, $\sum_{i=0}^{2}|a_i(0)| \neq 0$ then there exists $y(x)$ such that $\sum_{i=0}^{2} a_i(x)y(x)^i = 0$ and $y(x) = f(x) + O(x^K)$ where K is, at worst, $\frac{N}{2} + 1$. This is summarised in the following table.

	K
$D(0) \neq 0$	$N+2$
$D(x) = x^{2s}g(x)$ where $g(0) \neq 0$, $2s < N+1$, $a_0(x) \not\equiv 0$	$N+2-s$
$D(0) = 0$ and $a_0(x) \equiv 0$	$\min\{k \in \mathbf{N} : k \geq \frac{N}{2} + 1\}$
$D(x) \equiv 0$	$\min\{k \in \mathbf{N} : k \geq \frac{N}{2} + 1\}$

Table 1

It is natural to ask whether this type of result may be generalised to other Hermite-Padé approximations. This is an area where much work remains to be done.

In Chapters 4 and 5 attention was then turned to the practical problem of calculating $y(x)$ on some region in $\mathbf{C}$ and comparing this quadratic approximation with the appropriate Padé and Taylor approximations. The functions $\log(1+x), e^{-x}, \cos(x)$ and $\sqrt[3]{1+x}$ were chosen and

particular examples using these functions were studied in detail. In order to plot $y(x)$ several problems had to be solved. If $\sqrt{D(x)}$ denotes the principal square root then although

$$y(x) = \frac{-a_1(x) + \sqrt{D(x)}}{2a_2(x)}$$

or $$y(x) = \frac{-a_1(x) - \sqrt{D(x)}}{2a_2(x)}$$

in some neighbourhood of the origin, away from the origin it is not obvious which branch of the square root is the correct one. It is necessary to analytically trace a path from the origin to the point in question. A basic algorithm for doing this was given. Even then, there may be several analytic continuations on the region concerned, so in each case, one must restrict the region so that $y(x)$ is defined uniquely. This involves defining cuts from zeroes of $D(x)$.

Contour and surface plots of the quadratic, Padé and Taylor approximations and error functions were given over various regions. The quadratic approximation performs significantly better, over a larger region, in all the examples given.

Finally, in Chapter 6, three main topics were discussed. Firstly, it was shown that most sequences of quadratic Hermite-Padé forms of increasing order yield sequences of quadratic approximations of increasing order of accuracy. This seems to indicate that some sort of general convergence result may exist.

The question of multi-dimensional spaces of forms was then addressed. Given a multi-dimensional space of (A_2, A_1, A_0) forms it is of interest to know whether some choice of form from this space is better than another. It was shown that each linear space of (A_2, A_1, A_0) forms has a unique (up to constant factor) representative of maximal order. This, in fact, applies to all Hermite-Padé systems.

Lastly, attention was focussed on the structure of the table of quadratic Hermite-Padé forms. Much work has already been done on the structure of the Padé table and the introduction of the D-table was shown to provide an alternative proof of these results. This concept was generalised to give an idea of the types of structure which occur in more general Hermite-Padé tables.