On the Lambert W Function

by

R. M. Corless[1], G. H. Gonnet[2], D. E. G. Hare[3], D. J. Jeffrey[1], and D. E. Knuth[4]

[1]*Department of Applied Mathematics*
University of Western Ontario
London, Canada, N6A 5B7

[2]*Institut für Wissenschaftliches Rechnen*
ETH Zürich
Switzerland

[3]*Symbolic Computation Group*
University of Waterloo
Waterloo, Canada, N2L 3G1

[4]*Department of Computer Science*
Stanford University
Stanford, California
USA 94305-2140

Abstract

The Lambert W function is defined to be the multivalued inverse of the function $w \mapsto we^w$. It has many applications in pure and applied mathematics, some of which are briefly described here. We present a new discussion of the complex branches of W, an asymptotic expansion valid for all branches, an efficient numerical procedure for evaluating the function to arbitrary precision, and a method for the symbolic integration of expressions containing W.

1. Introduction

In 1758, Lambert solved the trinomial equation $x = q + x^m$ by giving a series development for x in powers of q. Later, he extended the series to give powers of x as well [48,49]. In [28], Euler transformed Lambert's equation into the more symmetrical form

$$x^\alpha - x^\beta = (\alpha - \beta)vx^{\alpha+\beta} \tag{1.1}$$

by substituting $x^{-\beta}$ for x and setting $m = \alpha\beta$ and $q = (\alpha - \beta)v$. Euler's version of Lambert's series solution was thus

$$\begin{aligned} x^n = 1 + nv &+ \tfrac{1}{2}n(n+\alpha+\beta)v^2 \\ &+ \tfrac{1}{6}n(n+\alpha+2\beta)(n+2\alpha+\beta)v^3 \\ &+ \tfrac{1}{24}n(n+\alpha+3\beta)(n+2\alpha+2\beta)(n+3\alpha+\beta)v^4 \\ &+ \text{etc.} \end{aligned} \tag{1.2}$$

After deriving the series, Euler looked at special cases, starting with $\alpha = \beta$. To see what this means in the original trinomial equation, we divide (1.1) by $(\alpha - \beta)$ and then let $\beta \to \alpha$ to get

$$\log x = vx^\alpha \ . \tag{1.3}$$

Euler noticed that if we can solve equation (1.3) for $\alpha = 1$, then we can solve it for any $\alpha \neq 0$. To see this, multiply equation (1.3) by α, simplify $\alpha \log x$ to $\log x^\alpha$, put $z = x^\alpha$ and $u = \alpha v$. We get $\log z = uz$, which is just equation (1.3) with $\alpha = 1$.

To solve this equation using (1.2), Euler first put $\alpha = \beta = 1$ and then rewrote (1.2) as a series for $(x^n - 1)/n$. Next he set $n = 0$ to obtain $\log x$ on the left-hand side and a nice series on the right-hand side:

$$\log x = v + \frac{2^1}{2!}v^2 + \frac{3^2}{3!}v^3 + \frac{4^3}{4!}v^4 + \frac{5^4}{5!}v^5 + \text{etc.} \tag{1.4}$$

This series, which can be seen to converge for $|v| < 1/e$, defines a function $T(v)$ called the *tree function* [41]. It equals $-W(-v)$, where $W(z)$ is defined to be the function satisfying

$$W(z)e^{W(z)} = z \ . \tag{1.5}$$

This paper discusses both W and T, concentrating on W.

The two functions are used in many applications: for example, in the enumeration of trees [10, 13, 67, 25, 41]; in the calculation of water-wave heights [64]; and in problems considered by Pólya and Szegö [56, Problem III.209, p. 146]. Wright used the complex branches of W, and roots of more general exponential polynomials, to solve linear constant-coefficient delay equations [69]. In [30], Fritsch, Shafer and Crowley presented an algorithm for the fixed-precision computation of one branch of $W(x)$ for $x > 0$. The computer algebra system Maple has had an arbitrary precision implementation of this same real-valued branch of W for many years, and since Release 2 has had an arbitrary precision implementation of all branches [14].

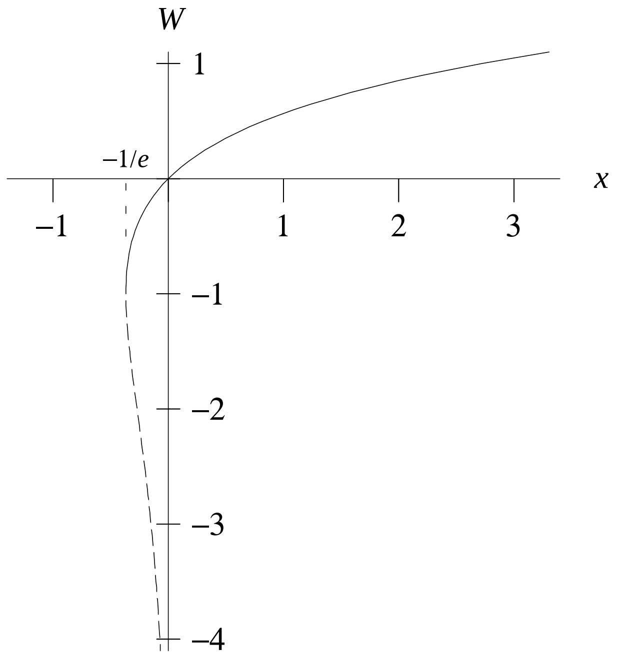

Figure 1. The two real branches of $W(x)$. ——, $W_0(x)$; − − −, $W_{-1}(x)$.

The purposes of this paper are to collect existing results on this function for convenient reference, and to present some new results on the asymptotics, complex analysis, numerical analysis, and symbolic calculus of the W function. We have examined the complex analytic properties of W and, building on the work of de Bruijn [11, pp. 27–28], determined the asymptotic expansions at both complex infinity and 0. This work led to an efficient and accurate implementation (in Maple V Release 2) of the arbitrary-precision complex floating-point evaluation of all branches of $W(z)$. Finally, some results of K. B. Ranger given at the Tenth Canadian Fluid Mechanics Symposium (also discussed in [58]) led us to rediscover a very old method for the integration of inverse functions, which allows the symbolic integration of a large class of functions containing W.

NOTATION

We use the letter W for this function, following early Maple usage. We propose to call it the *Lambert W function*, because it is the logarithm of a special case ($\beta = \alpha = -1$) of Lambert's series (1.2). Fortuitously, the letter W has additional significance because of the pioneering work on many aspects of W by E. M. Wright [69,70,71,72]. In [14], the function is also called the Omega function.

If x is real, then for $-1/e \le x < 0$ there are two possible real values of $W(x)$ (see Figure 1). We denote the branch satisfying $-1 \le W(x)$ by $W_0(x)$ or just $W(x)$ when there is no possibility for confusion, and the branch satisfying $W(x) \le -1$ by $W_{-1}(x)$. $W_0(x)$ is referred to as the *principal branch* of the W function. This notation will be explained and extended in section 4.

2. Applications

Since W is such a simple function, we would expect by Pareto's principle (eighty percent of your work is accomplished with twenty percent of your tools) that W would have many applications. In fact this is the case, although the presence of W often goes unrecognized. We present below a selection of applications.

Combinatorial applications

The tree function and related functions are used in the enumeration of trees. Let t_n be the number of rooted trees on n labelled points. The exponential generating function $T(x) = \sum t_n x^n/n!$ satisfies the functional equation $T(x) = x + xT(x) + xT(x)^2/2! + \cdots = xe^{T(x)}$, so $T(x)e^{-T(x)} = x$ and $T(x) = -W(-x)$. In [57], Pólya used this approach and Lagrange inversion to deduce that $t_n = n^{n-1}$, a formula that had previously been proved in other ways [10,13,67].

If we put

$$U(x) = T(x) - \tfrac{1}{2}T(x)^2 , \tag{2.1}$$

then one can show that $U(x)$ generates the number of labelled *unrooted* trees [24,41]. Similarly,

$$V(x) = \tfrac{1}{2}\log\frac{1}{1-T(x)} \tag{2.2}$$

generates the number of unicyclic components of a multigraph; subtract $\frac{1}{2}T(x) + \frac{1}{4}T(x)^2$ to generate the unicyclic components of a graph (when loops and parallel edges are not permitted). In fact, the exponential generating function for all connected graphs having a fixed excess of edges over vertices can be expressed as a function of $T(x)$ [72,41].

The number of mappings from $\{1, 2, \ldots, n\}$ into itself having exactly k component cycles is the coefficient of y^k in $t_n(y)$, where $t_n(y)$ is called the *tree polynomial* of order n (see [47]) and is generated by

$$\frac{1}{(1-T(z))^y} = \sum_{n\ge 0} t_n(y)\frac{z^n}{n!} . \tag{2.3}$$

One application of these functions is to derive the limiting distribution of cycles in random mappings [29]. Chaotic maps of the unit interval using floating-point arithmetic can be studied in this way; an elementary discussion that looks only at the expected length of the longest cycle can be found in [18].

Iterated exponentiation

The problem of iterated exponentiation is the evaluation of

$$h(z) = z^{z^{z^{z^{\cdot^{\cdot^{\cdot}}}}}} , \tag{2.4}$$

whenever it makes sense. Euler was the first to prove that this iteration converges for real z between e^{-e} and $e^{1/e}$. For convergence in the complex plane, it was shown in [5] that (2.4) converges for $\log z \in U = \{te^{-t} : |t| < 1 \text{ or } t^n = 1 \text{ for some } n = 1, 2, \ldots\}$ and that it diverges elsewhere. See [6] for further details and references.

It is not widely known, even though an old result (see [25]), that the function $h(z)$ has a closed-form expression in terms of T. When the iteration converges, it converges to

$$h(z) = \frac{T(\log z)}{\log z} = \frac{W(-\log z)}{-\log z}\,, \tag{2.5}$$

as can be seen on solving $h(z) = z^{h(z)}$ for $h(z)$ by taking logarithms. This immediately answers the question posed in [46] about the analytic continuation of $h(z)$.

Euler observed in [27] that the equation $g = z^{z^g}$ sometimes has real roots g that are not roots of $h = z^h$. A complete analysis of such questions, considering also the complex roots, involves the T function, as shown by Hayes in [37].

Solution of equations

The solution of $xe^x = a$ is $x = W(a)$ by definition, but Lémeray noted in [53] that a variety of other equations can be solved in terms of the same transcendental function. For example, the solution of $xb^x = a$ is $x = W(a \log b)/\log b$. The solution of $x^{x^a} = b$ is $\exp(W(a \log b)/a)$, and the solution of $a^x = x + b$ is $x = -b - W(-a^{-b} \log a)/\log a$.

Solution of a jet fuel problem

Consider the following equations, which describe the endurance E_t and range R of a jet airplane [3, *pp.* 312-323]. We wish to find the thrust specific fuel consumption c_t and the weight of the fuel $w_0 - w_1$, given the physical constants describing the plane and its environment. The equations are

$$E_t = \frac{C_L}{c_t C_D} \log\left(\frac{w_0}{w_1}\right)\,, \tag{2.6}$$

$$R = \frac{2}{c_t C_D}\left(\frac{2C_L}{\rho S}\right)^{1/2}\left(w_0^{1/2} - w_1^{1/2}\right)\,, \tag{2.7}$$

where E_t is the endurance, C_L and C_D are the lift and drag coefficients, w_0 is the initial weight of the plane, w_1 is the final weight of the plane, R is the range, S is the area of the horizontal projection of the plane's wings, and ρ is the ambient air density. We simplify these equations by first grouping the physical parameters and nondimensionalizing. We put $w = w_1/w_0$ and $c = E_t C_D c_t / C_L$ and introduce the new parameter

$$A = -\frac{\sqrt{2}E_t}{R}\left(\frac{w_0}{\rho S C_L}\right)^{1/2}\,. \tag{2.8}$$

The equations then become $c = -\log w$ and

$$2A\frac{1-\sqrt{w}}{\log w} = 1\,. \tag{2.9}$$

This equation has exactly one real solution if $A < 0$, since the left-hand side is a strictly increasing function of w. It can be solved in terms of W, the solution being

$$w = \begin{cases} A^{-2} W_0^2\left(Ae^A\right), & \text{if } A \le -1, \\ A^{-2} W_{-1}^2\left(Ae^A\right), & \text{if } -1 \le A < 0. \end{cases} \tag{2.10}$$

Once w is known, the thrust specific fuel consumption follows from $c = -\log w$.

Solution of a model combustion problem

The problem

$$\frac{dy}{dt} = y^2(1-y), \quad y(0) = \varepsilon > 0 \tag{2.11}$$

is used in [54,59] to explore perturbation methods. We show here that an explicit analytic solution is possible, in terms of W, and thus all the perturbation results in [54] can be simply tested by comparison with the exact solution. The model problem is separable and integration leads to an implicit form for the solution $y(t)$ (as given in [54]):

$$\frac{1}{y} + \log\left(\frac{1}{y} - 1\right) = \frac{1}{\varepsilon} + \log\left(\frac{1}{\varepsilon} - 1\right) - t\,, \tag{2.12}$$

which can be solved to get

$$y = \frac{1}{1 + W\left(ue^{u-t}\right)}\,. \tag{2.13}$$

where $u = 1/\varepsilon - 1$. Inspection of the differential equation shows that $0 < \varepsilon \le y < 1$ and this implies that the principal branch of W should be used. From this explicit solution, all the series results for this function in [54] can easily be verified.

Solution of an enzyme kinetics problem

In contrast to the last section, where W provided an exact solution to compare with perturbation expansions, this section gives an example where the perturbation expansion itself is done in terms of the W function.

In [54] the Michaelis–Menten model of enzyme kinetics is solved with a perturbation technique. A similar model, with a better scaling, is examined in [61]. The outer solution is taken to be of the form $s(\tau) = s_0(\tau) + \varepsilon s_1(\tau) + \ldots$ and $c(\tau) = c_0(\tau) + \varepsilon c_1(\tau) + \ldots$, and the leading order terms s_0 and c_0 are found to satisfy

$$c_0 = \frac{(\sigma+1)s_0}{\sigma s_0 + 1} \tag{2.14}$$

$$\frac{ds_0}{d\tau} = -\frac{(\sigma+1)s_0}{\sigma s_0 + 1}\,. \tag{2.15}$$

The equation for s_0 can be solved explicitly in terms of W.

$$s_0 = \frac{1}{\sigma} W\left(\sigma e^{\sigma-(\sigma+1)\tau}\right)\,. \tag{2.16}$$

This solution satisfies $s_0(0) = 1$ and uses the principal branch of the W function. The first order correction terms can also be expressed in terms of s_0, and hence in terms of W.

Solution of linear constant-coefficient delay equations

Perhaps the most significant use of the W function is in the solution of linear constant-coefficient delay equations [8,69]. Many of the complex-variable properties of this function (and generalizations of it) were proved by workers in this field, motivated by the appearance of W in the solution of simple delay equations such as the following:

$$\dot{y}(t) = ay(t-1) \tag{2.17}$$

subject to $y(t) = f(t)$, a known function, on $0 \le t \le 1$. This problem gains its significance from its occurrence in the study of the stability of nonlinear delay equations, for example.

One approach to solving this problem is to guess that $y = \exp(st)$ is a solution for some value of s. This gives immediately that

$$s\exp(st) = a\exp(st)\exp(-s) ,$$

or

$$s = W_k(a) ,$$

for any branch W_k. (See section 4 below for a description of the complex branches of W.) If $\exp(W_k(a)t)$ is a solution of $\dot{y} = ay(t-1)$, then by linearity so is

$$y = \sum_{k=-\infty}^{\infty} c_k \exp(W_k(a)t) , \tag{2.18}$$

for any 'reasonable' choice of c_k (i.e. such that the sum makes sense). One sees immediately that the solution will grow exponentially if any of the $W_k(a)$ have positive real part, and this leads to important stability theorems in the theory of delay equations [8].

This approach can be generalized to handle systems of constant-coefficient 'pure' delay equations, and scalar delay equations of the form

$$\dot{y} = ay(t-1) + by(t) . \tag{2.19}$$

We can also solve $\dot{y} = Ay(t-1) + By(t)$ where A and B are matrices, in the special case that A and B commute. Further generalizations require generalizations of the W function. See [8,19,69] for further discussion.

Resolution of a paradox in physics

We give here a brief overview of the recent use of W in [62,63] to explain an anomaly in the calculation of exchange forces between two nuclei within the hydrogen molecular ion (H_2^+). In considering the one-dimensional limit of this system, namely the double-well Dirac delta function model, the wave equation in atomic units

$$-\frac{1}{2}\frac{d^2\psi}{dx^2} - q\left[\delta(x) + \lambda\delta(x-R)\right]\psi = E(\lambda)\psi , \tag{2.20}$$

was used. The solution of this equation is expressed as a linear combination of *atomic orbital* solutions:

$$\psi = Ae^{-c|x|} + Be^{-c|x-R|}\,. \tag{2.21}$$

Making this solution continuous at each *well* ($x = 0$ and $x = R$) leads to the following transcendental equations for c:

$$c_\pm = q\left[1 \pm e^{-c_\pm R}\right]\,. \tag{2.22}$$

The solution of these equations can be expressed as

$$c_\pm = q + W(\pm qRe^{-qR})/R\,. \tag{2.23}$$

Scott *et al.* then go on to use this exact solution to explain how exponentially subdominant terms in the true asymptotic expansion of this solution account for differences between the predictions of previous numerical and asymptotic solutions for the model equations. In brief, the exponentially subdominant terms were missing from the previous asymptotic developments, but are still significant for small enough R.

Similarity solution for the Richards equation

Recent work uses both real branches of W to give a new exact solution for the Richards equation for water movement in soil [7]. By a similarity transformation, the Richards equation for the *moisture tension* Ψ,

$$\frac{d\theta}{d\Psi}\frac{\partial\Psi}{\partial t} = \frac{\partial}{\partial z}\left[K(\Psi)\frac{\partial\Psi}{\partial z} - K(\Psi)\right]\,, \tag{2.24}$$

is reduced, in a special case, to the ordinary differential equation

$$\alpha^* A\frac{dA}{dt} = 1 - A \tag{2.25}$$

which can be solved in terms of W as

$$A(t) = 1 + W\left[(-1 + A(0))\exp\left\{\frac{(A(0)-1)\alpha^* - t}{1 - A(0)}\right\}\right]\,. \tag{2.26}$$

Both real branches of W give rise to physically meaningful solutions. If we use W_0 the solution corresponds to *capillary rise*, while if instead we use W_{-1} the solution can be interpreted as *infiltration*.

Volterra equations for population growth

In [22, pp 102–109], we find an implicit analytic phase plane solution of the Volterra equations

$$\frac{dx}{dt} = ax(1-y), \quad \frac{dy}{dt} = -cy(1-x)\,, \tag{2.27}$$

essentially in terms of the W function (see equations (11) and (12) on page 104 of [22]). These equations, with $a = 2$ and $c = 1$, appear as problem B1 of the DETEST test suite for numerical methods for integration of ordinary differential equations [40]. The analytic solution is a closed loop in the phase plane. If the upper branch is y^+ and the lower y^-, then

$$\begin{aligned} y^+ &= -W_{-1}\left(-Cx^{-c/a}e^{cx/a}\right) , \\ y^- &= -W_0\left(-Cx^{-c/a}e^{cx/a}\right) , \end{aligned} \tag{2.28}$$

where C is an arbitrary constant.

The remaining problem is to find an expression for t, and we can find this in terms of quadrature:

$$t = \int_{x_0}^{x} \frac{d\xi}{a\xi(1-y(\xi))} = \int_{y_0}^{y} \frac{d\eta}{-c\eta(1-x(\eta))} . \tag{2.29}$$

There are square-root singularities and branches in these integrals, but these can be handled with standard changes of variables.

Asymptotic roots of trinomials

The sequence of polynomial equations

$$(a+n)x^n + (b-n)x^{n-1} + f(n) = 0 \tag{2.30}$$

has real roots near 1 having an asymptotic series

$$x_n = 1 + \frac{y_n - a - b}{n} + O(y_n^2 n^{-2}) , \tag{2.31}$$

where

$$y_n = W(-e^{a+b}f(n)) . \tag{2.32}$$

See [33, page 215], where a more detailed formula correct to $O(y_n^3 n^{-3})$ is given.

Epidemics and components

Suppose everyone in a population of n people is in close contact with α others, chosen at random. If one person is infected with a disease and if the disease spreads by transitivity to all those who are in close contact with the infected person, the total number of infected people will be approximately γn for large n, where

$$\gamma = 1 - e^{-\alpha\gamma} . \tag{2.33}$$

This formula, derived heuristically for fixed integer α by Solomonoff and Rapoport [65], then proved rigorously by Landau [50], holds also when α is an expected value (not fixed for all individuals, and not necessarily an integer). Since (2.33) can be written

$$\alpha e^{-\alpha} = \alpha(1-\gamma)e^{\alpha(\gamma-1)} ,$$

we have

$$\gamma = 1 - T(\alpha e^{-\alpha})/\alpha = 1 + W(-\alpha e^{-\alpha})/\alpha \tag{2.34}$$

when $\alpha \geq 1$, using the principal branches of T and W.

The epidemic or reachability problem is closely related to the size of the 'giant component' in a random graph, a phenomenon first demonstrated in a famous paper by Erdős and Rényi [26]. When a graph on n vertices has $m = \frac{1}{2}\alpha n$ edges chosen at random, for $\alpha > 1$, it almost surely has a connected component with approximately γn vertices, where γ is given by (2.34). The study of the emergence of this giant component when $\alpha = 1 + \beta n^{-1/3}$ is particularly interesting [41].

Analysis of algorithms

The behaviour of epidemics, random graphs and other dynamic models gives insight also into the behaviour of computer methods and data structures, so it is not surprising that the T and W functions occur frequently in theoretical computer science. For example, algorithms that are based on a divide-and-conquer paradigm related to random graphs require the analysis of solutions to the recurrence relation

$$x_n = c_n + \frac{1}{n-1}\sum_{k=1}^{n-1}\binom{n}{k}\left(\frac{k}{n}\right)^{k-1}\left(\frac{n-k}{n}\right)^{n-k-1} x_k \tag{2.35}$$

for various given sequences $c_0, c_1, \ldots$ and the theory of this recurrence depends on the behaviour of $W(x)$ near its quadratic singularity at $x = -1/e$ [47]. Many derivations in algorithmic analysis depend on generating functions, and the formulae

$$\left(\frac{T(z)}{z}\right)^a = e^{aT(z)} = \sum_{n=0}^{\infty} a(a+n)^{n-1}\frac{z^n}{n!}\ , \tag{2.36}$$

$$\frac{e^{aT(z)}}{1-T(z)} = \sum_{n=0}^{\infty}(a+n)^n\frac{z^n}{n!} \tag{2.37}$$

have proved to be especially useful. We also have

$$[z^n]F\big(T(z)\big) = [t^n]F(t)(1-t)e^{nt}\ , \tag{2.38}$$

$$[z^n]F\big(W(z)\big) = [t^n]F(t)(1+t)e^{-nt}\ , \tag{2.39}$$

for any power series F, where $[z^n]$ extracts the coefficient of z^n. These formulae are consequences of the Lagrange inversion theorem.

One of the most important methods for information retrieval is a technique known as *hashing with uniform probing*: each of n items is mapped into a random permutation $(p_1, \ldots p_m)$ of $\{1, \ldots, m\}$ and stored in the first cell p_j that is currently unoccupied. Gonnet proved in [34] that the expected maximum number of probes (the maximum j) over all n items, when $n = \alpha m$ and $\alpha < 1$, is approximately $T(m \log_2 \alpha)/\log \alpha - 1$ for large m.

Another, quite different, application to information retrieval concerns the expected height of random binary search trees [23,60]. Let binary search trees with n nodes be constructed by standard insertions from a random permutation of $1, \ldots, n$; let h_n be a random variable giving the height of such trees. Devroye proved in [23] that the expected value $E(h_n)$ obeys

$$\lim_{n\to\infty}\frac{E(h_n)}{\log n} = c = \frac{1}{T(1/(2e))}\ , \tag{2.40}$$

and, moreover, for all $\epsilon > 0$,

$$\lim_{n\to\infty} P(h_n \ge (c+\epsilon)\log n) = 0 \quad \text{and} \quad \lim_{n\to\infty} P(h_n \le (c-\epsilon)\log n) = 0\ .$$

Thus $h_n \to c\log n$ in probability as $n \to \infty$.

Pedagogical applications

The W function provides a useful exercise for first-year calculus students in defining implicit functions, differentiating such functions, computing Taylor series, and, as we shall see, computing antiderivatives. For introductory numerical analysis classes, it provides a good root-finding problem (suitable for Newton's method or Halley's method), an example of a well-conditioned function for evaluation (away from the branch point $x = -1/e$) and an ill-conditioned function for evaluation (near the branch point). For complex variable courses, it provides a useful example, similar to the logarithm and the square root, of a multivalued function. In fact, it is probably the simplest function that exhibits both algebraic and logarithmic singularities. It also provides a simple example of the application of the Lagrange inversion theorem [12]. The asymptotic analysis of the W function might profitably be used in a later course.

3. Calculus

The principal branch of W is analytic at 0. This follows from the Lagrange inversion theorem (see e.g. [12]), which gives the series expansion below for $W_0(z)$:

$$W_0(z) = \sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} z^n \,. \tag{3.1}$$

It is an interesting exercise to calculate this series by directly reverting the series for we^w. Because of the derivative singularity of W_0 at $z = -1/e$ (the existence of which follows from the derivative of $w \mapsto we^w$ being 0 at $w = -1$), the radius of convergence of the series (3.1) must be less than or equal to $1/e$. That it is in fact exactly equal to $1/e$ is easily seen using the ratio test.

Speaking of convergence, Euler did not mention that the series (1.2) converges at least for

$$M_E(\alpha, \beta)|v| < \frac{1}{e} \,,$$

where $M_E(a,b) = \sqrt{(a^2+b^2)/2}$ is the 'Euclidean mean' of a and b. This result follows, for example, by applying the ratio test and then converting α and β to polar coordinates. For most α and β the series will converge for larger v, but if $\alpha = \beta$ this is exactly the radius of convergence of the series. The details of this calculation, which are not completely trivial, are left as an exercise for the reader. Since the series for W can be derived from (1.2) by putting $\alpha = \beta = -1$ we see that this result implies the convergence result for W above.

Differentiating the defining equation $x = W(x)e^{W(x)}$ for W and solving for W', we obtain the following expressions for the derivative of W:

$$\begin{aligned} W'(x) &= \frac{1}{(1+W(x))\exp(W(x))} \\ &= \frac{W(x)}{x(1+W(x))} \,, \qquad \text{if } x \neq 0. \end{aligned} \tag{3.2}$$

Historically, computer algebra systems have been quite cavalier about the handling of exceptional points. The equation above is a typical example, for Maple V Release 3 will return $W(x)/(x(1+W(x)))$ when asked to differentiate W, and hence is able to compute $W_0'(0) = 1$ only as a limit. See [20] for further discussion of the handling of special cases (the so-called specialization problem) by computer algebra systems.

Taking further derivatives, we can see by induction that the nth derivative of W is

$$\frac{d^n W(x)}{dx^n} = \frac{\exp(-nW(x))p_n(W(x))}{(1+W(x))^{2n-1}} \qquad \text{for } n \geq 1\,, \tag{3.3}$$

where the polynomials $p_n(w)$ satisfy the recurrence relation

$$p_{n+1}(w) = -(nw + 3n - 1)p_n(w) + (1+w)p_n'(w)\,, \qquad \text{for } n \geq 1\,. \tag{3.4}$$

The initial polynomial is $p_1(w) = 1$. The value of $p_n(0)$ is $(-n)^{n-1}/n!$ for $n \geq 1$. Although the polynomials p_n do not seem to be known in other contexts, there is a similar formula

$$\frac{d^n W(e^x)}{dx^n} = \frac{q_n\big(W(e^x)\big)}{\big(1+W(e^x)\big)^{2n-1}} \qquad \text{for } n \geq 1\,, \tag{3.5}$$

in which the polynomials q_n, given by

$$q_n(w) = \sum_{k=0}^{n-1} \left\langle\!\!\left\langle {n-1 \atop k} \right\rangle\!\!\right\rangle (-1)^k w^{k+1}\,, \tag{3.6}$$

contain coefficients expressed in terms of the second-order Eulerian numbers [35]. We have

$$q_{n+1}(w) = -(2n-1)wq_n(w) + (w + w^2)q_n'(w) \tag{3.7}$$

and $q_1(w) = w$.

Taking logarithms of both sides of $We^W = z$ and rearranging terms, we obtain the simplification transformation

$$\log W(z) = \log z - W(z)\,, \tag{3.8}$$

which is valid at least for the principal branch when $z > 0$. See [43] for a more general formula.

We turn now to the question of integrating expressions containing the W function. In [58], K. B. Ranger used the following example to illustrate some attempts to integrate the Navier-Stokes equations in parametric form:

$$x = pe^p\,, \tag{3.9}$$

$$\frac{dy}{dx} = p\,. \tag{3.10}$$

Differentiating equation (3.9) with respect to y gives

$$\frac{dx}{dy} = \frac{dp}{dy}e^p + pe^p\frac{dp}{dy}\,, \tag{3.11}$$

and simplifying using equation (3.10) we then obtain

$$\frac{dy}{dp} = p(p+1)e^p \ , \tag{3.12}$$

which is easily integrated to give

$$y = (p^2 - p + 1)e^p + C \ . \tag{3.13}$$

Since y is clearly an antiderivative of $W(x)$, Ranger had discovered a simple technique for symbolic integration of $W(x)$. In our notation,

$$\begin{aligned} \int W(x)\,dx &= (W^2(x) - W(x) + 1)e^{W(x)} + C \\ &= x(W(x) - 1 + 1/W(x)) + C \ . \end{aligned} \tag{3.14}$$

When one tries this technique on other functions containing W, one sees that it is really a special change of variable. Putting $w = W(x)$, so $x = we^w$ and $dx = (w+1)e^w\,dw$, we see for example that

$$\begin{aligned} \int xW(x)\,dx &= \int we^w \cdot w \cdot (1+w)e^w\,dw \\ &= \tfrac{1}{8}(2w-1)(2w^2+1)e^{2w} + C \\ &= \tfrac{1}{2}\left(W(x) - \tfrac{1}{2}\right)\left(W^2(x) + \tfrac{1}{2}\right)e^{2W(x)} + C \ . \end{aligned} \tag{3.15}$$

This is valid for all branches of W, as by definition $\frac{d}{dw}we^w \neq 0$ at any interior point of any branch.

The problem of integrating expressions containing W is a special case of integrating expressions containing inverse functions. By using the technique described above, several authors have rediscovered the following formula [55].

$$\int f^{-1}(x)\,dx = yf(y) - \int f(y)\,dy \ . \tag{3.16}$$

Finally, note that this technique allows the Risch algorithm to be applied to determine whether integrals containing W are elementary or not. For an introduction to the Risch algorithm see for example [31].

4. Branches and Asymptotics

We have seen that W has two branches on the real line. We have also seen, in the delay equations example, that complex values for W are required. Thus, to extend W to the complex plane, we must define all of the branches of $W(z)$. For the sake of clarity, the following discussion is quite detailed. We begin by recalling the definition of the branches

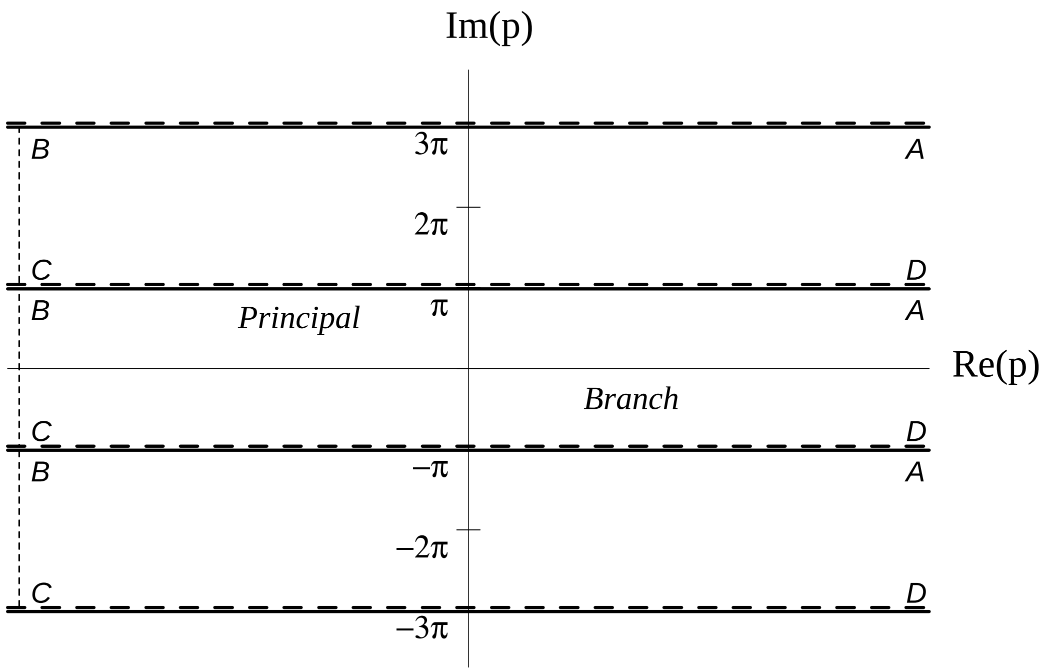

Figure 2. The ranges of the branches of $p = \log z$. The range of the principal branch is $-\pi < \mathrm{Im}(p) \le \pi$. The pairs of heavy lines, one solid, one dashed, together show one boundary of the branch. The solid line shows that the branch is closed when the boundary is approached from below. The points A, B, C, and D are mapped by $z = e^p$ onto the corresponding points in Figure 3. The pair of heavy lines in Figure 3 correspond not to a pair from this figure, but to the two lines on either side of a branch.

---

of the complex logarithm, partly to establish our notation and partly because we use the complex logarithm later.

If $p = \log z$, then $z = e^p$. We follow standard usage and say that z is in the z-plane and p is in the p-plane. One set of branches for the logarithm is obtained by partitioning the p-plane with horizontal boundary lines at $p = (2k+1)\pi i$, as shown in Figure 2. Each of the regions so defined then maps precisely onto the z-plane minus $(-\infty, 0]$. Further, it is nearly universal to consider the points on the boundary between two regions as belonging to the region below them. In other words, the boundary is *attached* to the region below it. This is shown on the figure by drawing two lines for each boundary: a solid line showing the points on the boundary attached to the region below them, and a dashed line showing points next to the boundary but belonging to the region above them. The region straddling the origin ($-\pi < \mathrm{Im}(p) \le \pi$) defines the range of the principal branch of the logarithm.

The z-plane corresponding to Figure 2 is shown in Figure 3. All of the solid boundary lines in Figure 2 map onto the solid line running along the negative real axis, and all of the dashed near-boundary lines map onto the dashed line just below the axis. The letters A, B, C, D further indicate the geometry of the map.

The negative real axis in the z-plane is called the branch cut for the logarithm, and the limiting value $z = 0$ is called the branch point. The branch choices shown in Figures 2 and 3

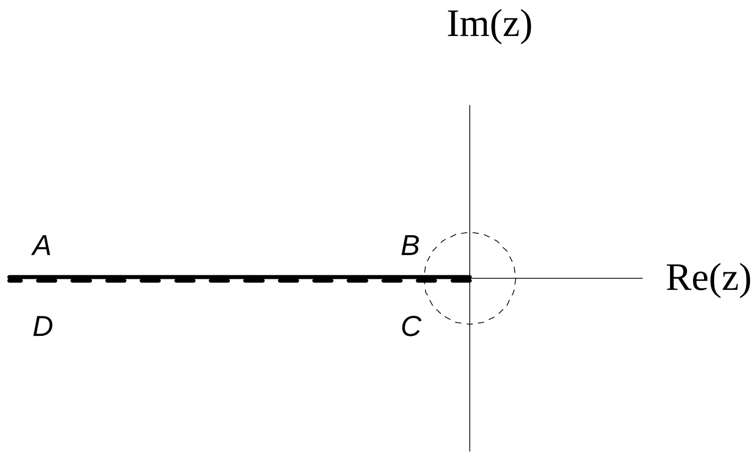

Figure 3. The branch cut for $p = \log z$. The heavy solid line is the image under $z = e^p$ of the solid lines in Figure 2. The dashed line from C to D is the image of the similarly dashed lines in Figure 2, indicating in this figure the open edge of the domain of $p = \log z$. The dashed circle running counterclockwise from C to B is the image of the similarly dashed lines in Figure 2, with closure at B.

---

conform to the rule of *counter-clockwise continuity* (CCC) around the branch point, which is a mnemonic principle that gives some uniformity to choices for the branch cuts of all the elementary functions [44]. Here, this convention distinguishes between two possibilities, namely the choice of attaching the image of the boundary in the p-plane to the top or to the bottom of the branch cut in the z-plane. The phrase *counter-clockwise continuous* is intended to convey the idea that a curve is being drawn around the branch point $z = 0$, its start and end points being on the branch cut and distinguished by the counter-clockwise sense. The image of any such curve is continuous in the limit as we approach its end, meaning in this case the positive side of the branch cut. In Figure 3, the curve is a circle, which we traverse from C to B, and then CCC ensures that the image of this circle is closed at the image of B in the p-plane.

Turning to the W function, we put $w = W(z)$ and $z = we^w$. We shall now specify the boundary curves that maximally partition the w-plane and find the images of these boundary curves, which are the branch cuts in the z-plane. If $w = \xi + i\eta$ and $z = x + iy$, then

$$x = e^{\xi}\left(\xi\cos\eta - \eta\sin\eta\right)\,, \tag{4.1}$$

$$y = e^{\xi}\left(\eta\cos\eta + \xi\sin\eta\right)\,. \tag{4.2}$$

We should like the z-plane branch cut(s) for W to be similar to that for the logarithm, and therefore we consider the images of the negative real z-axis in the w-plane. If $y = 0$ in (4.2) then

$$\eta = 0 \quad \text{or} \quad \xi = -\eta\cot\eta\,. \tag{4.3}$$

If further $\xi\cos\eta - \eta\sin\eta < 0$, then the images are precisely the *negative* real z-axis. See

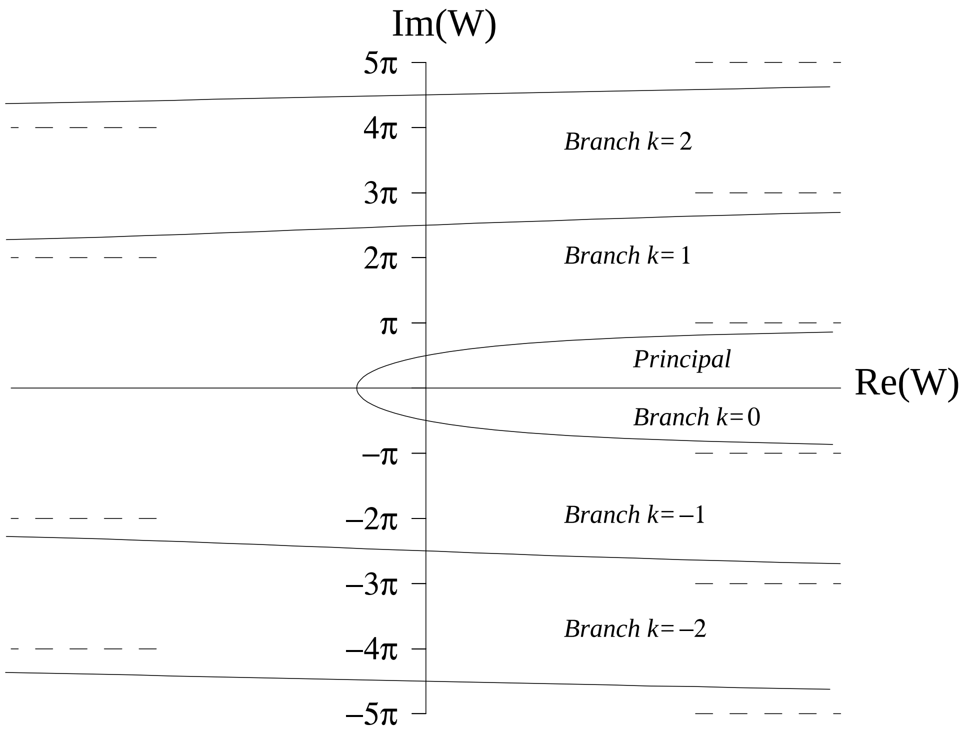

Figure 4. The ranges of the branches of $W(z)$. This figure does not contain closure information, which is given in the separate detailed figures of individual branches. Each branch is given a number, the principal branch being numbered 0. The boundaries of the branches are asymptotic to the dashed lines, which are horizontal at multiples of π.

Figure 4. We number the resulting regions in the w-plane as indicated in the figure, and denote the branch of W taking values in region k by $W_k(z)$.

The curve which separates the principal branch, W_0, from the branches W_1 and W_{-1} is

$$\{-\eta \cot \eta + \eta i : -\pi < \eta < \pi\} \tag{4.4}$$

together with -1 (which is the limiting value at $\eta = 0$). The curve separating W_1 and W_{-1} is simply $(-\infty, -1]$. Finally, the curves separating the remaining branches are

$$\{-\eta \cot \eta + \eta i : 2k\pi < \pm\eta < (2k+1)\pi\} \quad \text{for } k = 1, 2, \ldots \tag{4.5}$$

These curves, the inverse images of the negative real axis under the map $w \mapsto we^w$, partition the w-plane. It needs to be shown that each partition maps bijectively onto the z-plane. This can be established by an appeal to the Argument principle, or more simply by noting that the Jacobian of the transformation (considered as a map from $\mathbf{R}^2 \to \mathbf{R}^2$) is $e^{2\xi}\left((\xi+1)^2+\eta^2\right)$, which is nonzero everywhere except at the branch point. This implies by the inverse function theorem that simple curves surrounding the branch point, approaching

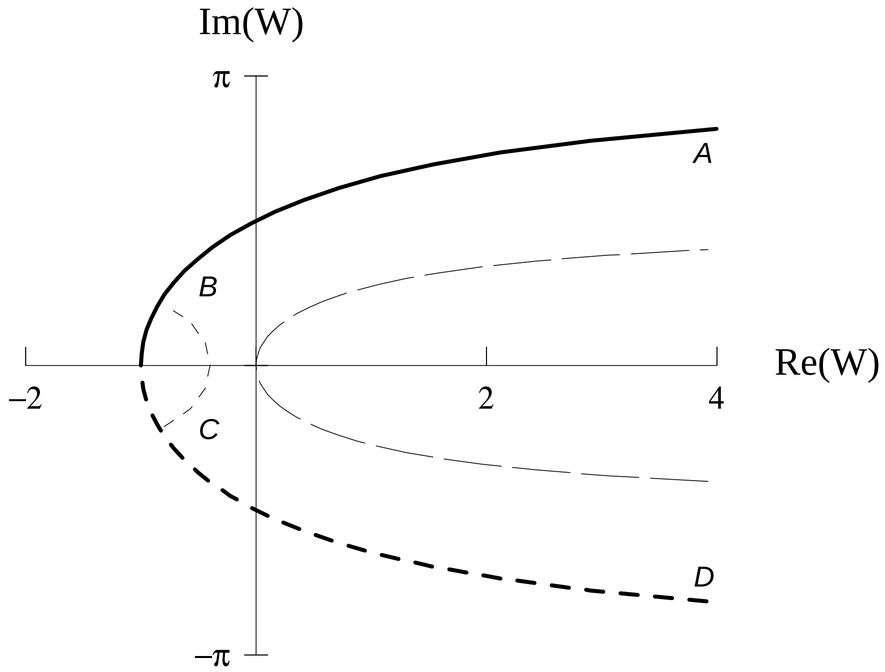

Figure 5. The range of the principal branch, $W_0(z)$. The heavy solid line again indicates closure. The points A, B, C, and D are the images of the corresponding points in Figure 6. The light dashed line — — — to the right of the imaginary axis is the image of the imaginary axis in Figure 6.

---

the branch cut from opposite sides, map to curves from one inverse image of the branch cut to another in the w-plane.

Remark. The curves (4.4) and (4.5) together form a subset of the so-called *Quadratrix of Hippias* [9], the missing parts being the curves corresponding to $(2k-1)\pi < \pm\eta < 2k\pi$. That is, the Quadratrix of Hippias is the union of the images of the real axis under the various branches of W, excluding $(-\infty, -1)$.

$W_0(z)$ is special, as it is the only branch which contains any part of the positive real axis in its range, and as noted previously we call this the principal branch of $W(z)$. It has a second-order branch point at $z = -1/e$ corresponding to $w = -1$, which it shares with both $W_{-1}(z)$ and $W_1(z)$. $W_0(z)$ is analytic at $z = 0$, with value $W_0(0) = 0$. Its branch cut is $\{z : -\infty < z \leq -1/e\}$. We choose to close the branch cut on the top, so $W_0(z)$ has counter-clockwise continuity (CCC) around the branch point $z = -1/e$. Thus the image of the curve $z = -1/e + \varepsilon e^{it}$ around the branch point in the z-plane, for $-\pi < t \leq \pi$, is a continuous curve in the region labelled 0. See Figure 5 and Figure 6.

Because of the extra branch point, W_{-1} and W_1 each have a double branch cut, $\{z : -\infty < z \leq -1/e\}$ and $\{z : -\infty < z \leq 0\}$. We close the branch cuts as before on the top. The function is thus CCC on the cut from the branch point at $z = 0$. This choice of closure implies that $W_{-1}(z)$ is real for $z \in [-1/e, 0)$. Thus $W_0(z)$ and $W_{-1}(z)$ are the only branches of W that take on real values. See Figure 7, Figure 8, and Figure 9. All

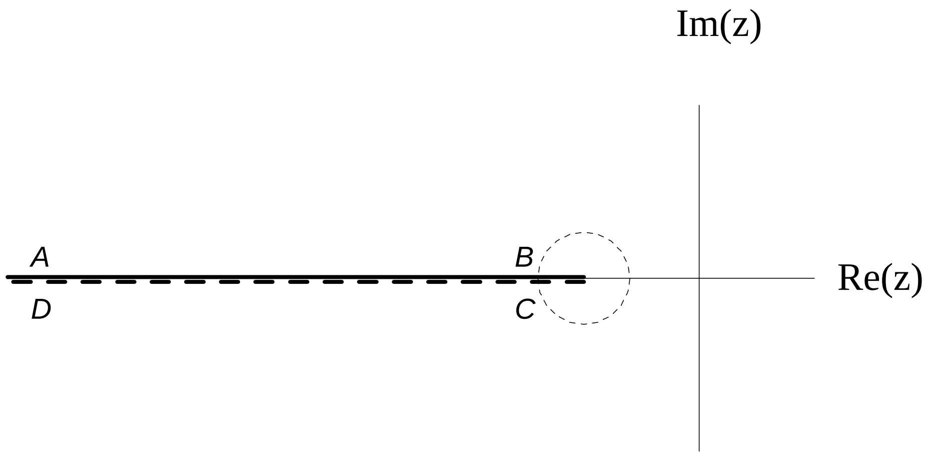

Figure 6. The branch cut for $W_0(z)$. The heavy solid line is the image under $z = we^w$ of the heavy solid line in Figure 5, and similarly for the dashed line. The dashed circle running counterclockwise from C to B is the image of the similarly dashed line in Figure 5, and is closed at B. The branch cut is $\{z : -\infty < z \le -1/e\}$.

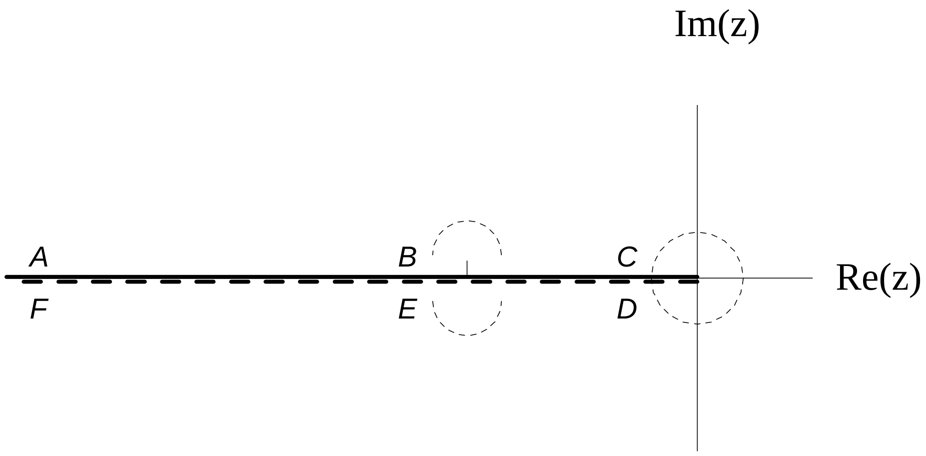

Figure 7. The branch cut for $W_k(z)$, $k \neq 0$. Closure is indicated by a heavy solid line. For $W_1(z)$ and $W_{-1}(z)$, the dashed semicircles centred at $z = -1/e$ are the images under $z = we^w$ of the corresponding arcs in Figures 8 and 9. The dashed circle running counter-clockwise from D to C, closed at C, is the image of a line running from the corresponding points D to C in each of Figures 8 and 9.

---

other branches of W have only the branch cut $\{z : -\infty < z \le 0\}$, closed on the top for counter-clockwise continuity, and thus are similar to the branches of the logarithm.

For all multivalued functions, the division of the complex plane into branches is somewhat arbitrary, and even the elementary functions do not have universally accepted

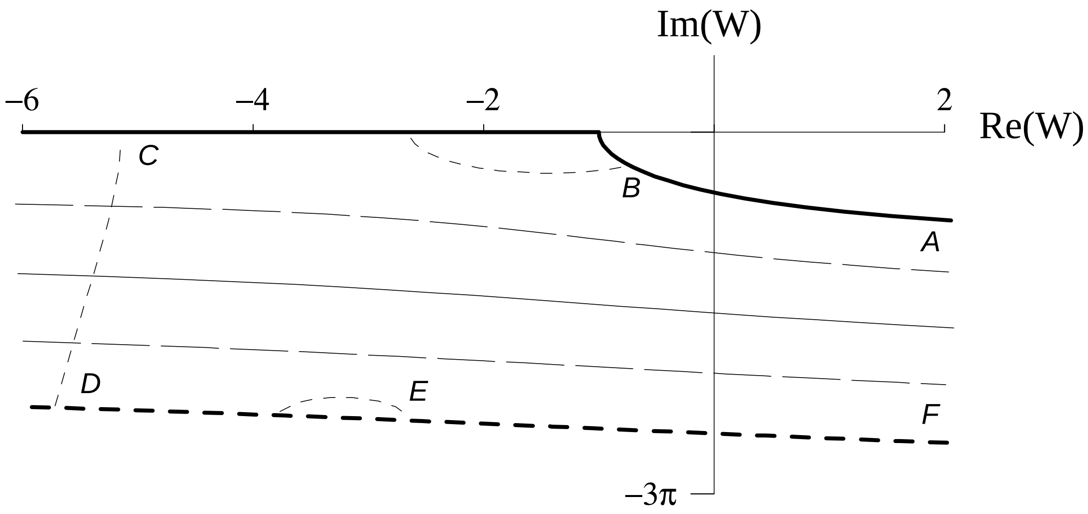

Figure 8. Details of the range of $W_{-1}(z)$. The curve from positive infinity, through A, to the 'corner' at $W = -1$ is the image of $z = (-\infty, -1/e]$. The curve from $W = -1$, through C to negative infinity is the image of $z = [-1/e, 0)$. Thus, the range of $W_{-1}(z)$ includes part of the real line by this choice of closure. The light solid line and the light dashed lines that cut CD and the imaginary axis are images of the positive real axis and the imaginary axis respectively.

---

branches [44]. The benefits of our choices for the Lambert W function are as follows. Most importantly, the coincidence of the branch cut for the branch point at 0 with the corresponding branch cut for the logarithm and the fact that both functions are CCC yield nice asymptotic expansions of the branches of W at both 0 and (complex) infinity. Secondly, the placement of the branch cuts in the real axis implies that W has *near conjugate symmetry*, meaning that except for points in the branch cut we have $W_k(\overline{z}) = \overline{W_{-k}(z)}$ for all integers k. Systems having a signed 0 (for example, ones implementing *IEEE Standard 754* [2]) can exploit this symmetry on the branch cut as well.

The main drawback with this choice of branches is that one particular expansion below, namely equation (4.22), is more complicated than it would be if the branch cuts were chosen as $\{z : -\infty < z \le -1/e\}$ and $\{z : 0 \le z < \infty\}$.

Asymptotic expansions

To develop the asymptotic expansions of the branches of W at both 0 and (complex) infinity, we note that in both cases it is the exponential character of $w \mapsto we^w$ which dominates, and hence the leading term of such an expansion will be some form of logarithm. (Unless, of course, the principal branch at $z = 0$ is being considered.) We write

$$w = W(z) = \log z + u \, , \tag{4.6}$$

where for z either close to 0 or large we expect u to be small relative to $\log z$, and where we are not specifying at this point which branch of the logarithm we intend to use. Sub-

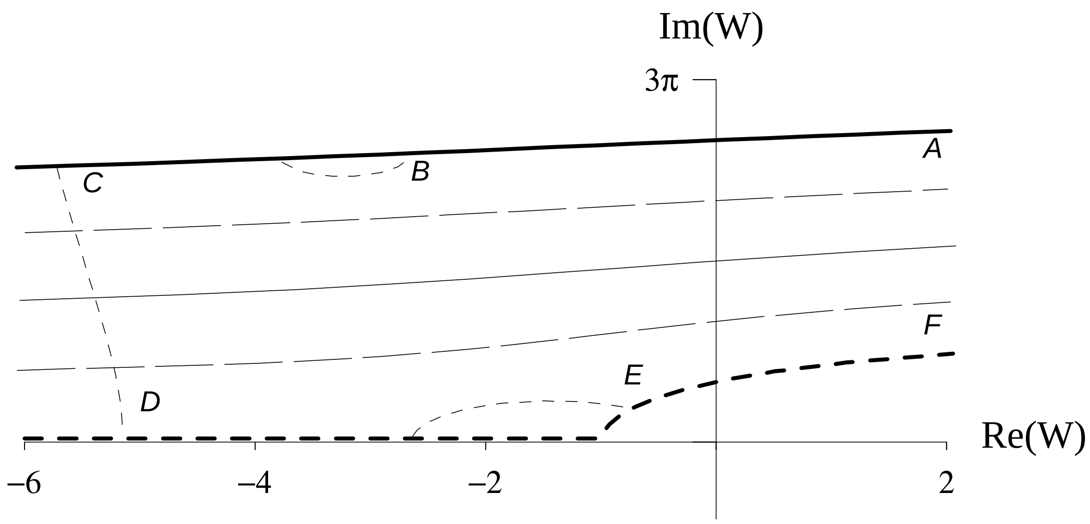

Figure 9. Details of the range of $W_1(z)$. The curve passing through A and C is the image of the negative real z-axis. Notice that $W_1(z)$ contains no part of the real axis in its range, by choice of closure. The light solid line is the image of the positive real axis, and the light dashed lines are the image of the imaginary axis.

---

stituting this expression into the defining relation for W, namely $we^w = z$, we obtain

$$(\log z + u)ze^u = z \, , \tag{4.7}$$

and hence

$$(1 + \frac{u}{\log z})e^u = \frac{1}{\log z} \, . \tag{4.8}$$

Under the assumption that $|u| \ll |\log z|$ we have $e^u \sim 1/\log z$, hence

$$u \sim \log \frac{1}{\log z} \, , \tag{4.9}$$

where it must be stressed that the two logarithms in (4.9) need not be the same. However, the size assumption on u is best satisfied by choosing the principal branch for the outer logarithm in (4.9), so we will make that choice at this point. To emphasize that the branch of the inner logarithm is not yet chosen, we will rewrite (4.9) using Log to denote the inner logarithm:

$$u \sim \log \frac{1}{\mathrm{Log}\, z} \, . \tag{4.10}$$

One further observation to make here is that so long as $\mathrm{Log}\, z$ is not a negative real number we can replace $\log \frac{1}{\mathrm{Log}\, z}$ with $-\log \mathrm{Log}\, z$. $\mathrm{Log}\, z$ can only be a negative real number if Log is chosen to be the principal branch of the logarithm and z is a positive real number less than 1. Keeping this in mind, then, we now adapt de Bruijn's argument [11, pp. 27–28] establishing the asymptotic expansions for the branches of W at 0 and infinity. (de Bruijn computes the expansion only for real $z \to \infty$, but we will see below that the method is valid for all branches and for $z \to 0$ as well. de Bruijn comments that the proof is modelled on the usual proof of the Lagrange Inversion Theorem.) It is interesting that the asymptotic series are, in fact, convergent.

Continuing from (4.6) and (4.10), write

$$w = \operatorname{Log} z - \log \operatorname{Log} z + v\,. \tag{4.11}$$

On substituting (4.11) into $we^w = z$ we get

$$\frac{(\operatorname{Log} z - \log \operatorname{Log} z + v)e^v z}{\operatorname{Log} z} = z\,. \tag{4.12}$$

For convenience denote $1/\operatorname{Log} z$ by σ and $\log \operatorname{Log} z/\operatorname{Log} z$ by τ, as de Bruijn does. Then, assuming $z \neq 0$ and $\operatorname{Log} z \neq 0$,

$$e^{-v} - 1 - \sigma v + \tau = 0\,. \tag{4.13}$$

This is equation (2.4.6) in de Bruijn, except that we are interpreting the logarithms as being possibly any branch, and have performed only simplifications valid for all branches of the logarithm. We now ignore the relation that exists between σ and τ, and consider them as small independent complex parameters.

We will show that there exist positive numbers a and b such that if $|\sigma| < a$ and $|\tau| < a$ then equation (4.13) has exactly one solution in the domain $|v| < b$, and that this solution is an analytic function of both σ and τ in the region $|\sigma| < a$, $|\tau| < a$.

Let δ be the lower bound of $|e^{-\zeta} - 1|$ on the circle $|\zeta| = \pi$. (There is nothing special about the number π, here—any number less than 2π would do. However, this is the number de Bruijn chose.) Then δ is positive, and $e^{-\zeta} - 1$ has just one root inside that circle, that is, at $\zeta = 0$. If we now choose the positive number $a = \delta/(2(\pi + 1))$, then we have

$$|\sigma\zeta - \tau| \leq |\sigma||\zeta| + |\tau| < \tfrac{1}{2}\delta \tag{4.14}$$

for $|\sigma| < a, |\tau| < a$ and $|\zeta| = \pi$. This means that $|e^{-\zeta} - 1| > |\sigma\zeta - \tau|$ on the circle $|\zeta| = \pi$. Thus, by Rouché's Theorem [68, §3.42], equation (4.13) has exactly one root, v, inside the circle. By Cauchy's Theorem,

$$v = \frac{1}{2\pi i}\int_{|\zeta|=\pi} \frac{-e^{-\zeta} - \sigma}{e^{-\zeta} - 1 - \sigma\zeta + \tau}\zeta\, d\zeta\,, \tag{4.15}$$

where the integration contour is taken in the counterclockwise direction.

Remark This already establishes that for any branch choice of the logarithm denoted by Log in (4.11) (which all differ by multiples of $2\pi i$) there is exactly one number $v \in |\zeta| < \pi$ so that w defined by (4.11) satisfies $we^w = z$. This establishes the existence of the infinite number of roots $W_k(z)$ (at least for large z and for z near zero). This point was not noted in [11].

Now for every ζ on the integration path, $|\sigma\zeta| + |\tau|$ is less than $\frac{1}{2}|e^{-\zeta} - 1|$, so we can expand the denominator of the integrand of (4.15) in an absolutely and uniformly convergent power series

$$\frac{1}{e^{-\zeta} - 1 - \sigma\zeta + \tau} = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} (e^{-\zeta} - 1)^{-k-m-1} \zeta^k \sigma^k \tau^m (-1)^m C_m^{m+k} . \tag{4.16}$$

Substituting (4.16) into (4.15) and integrating term by term, we obtain v as the sum of an absolutely convergent double power series in σ and τ. Notice that all terms not containing τ vanish, because the corresponding integrands are regular at $\zeta = 0$. Therefore, we can conclude that (4.13) has exactly one solution v satisfying $|v| < \pi$, and this solution can be written as

$$v = \sum_{k=0}^{\infty} \sum_{m=1}^{\infty} c_{km} \sigma^k \tau^m , \tag{4.17}$$

where the c_{km} are constants not depending on σ or τ.

Now return to the special values of σ and τ as functions of z. For z sufficiently large, we have $|\sigma| < a$ and $|\tau| < a$, and likewise for z sufficiently small (this point was not noted in [11]). Thus we have established that the following formula gives the asymptotics for *all* non-principal branches of W both at infinity and at 0:

$$W(z) = \operatorname{Log} z - \log \operatorname{Log} z + \sum_{k=0}^{\infty} \sum_{m=1}^{\infty} c_{km} (\log \operatorname{Log} z)^m (\operatorname{Log} z)^{-k-m} . \tag{4.18}$$

The coefficients c_{km} can be found using the Lagrange Inversion Theorem [12]. In [15, pp 228–229], Comtet observed that solving (4.13) for v in terms of τ is equivalent to finding the inverse of the function $1 - e^{-v} + \sigma v$, and thus obtained $c_{km} = \frac{1}{m!}(-1)^k \left[{k+m \atop k+1}\right]$, where $\left[{k+m \atop k+1}\right]$ is a Stirling cycle number [35,42].

The series in (4.18), being absolutely convergent, can be rearranged into the form

$$\begin{aligned} W(z) =& L_1 - L_2 + \frac{L_2}{L_1} + \frac{L_2\,(-2 + L_2)}{2\,{L_1}^2} + \frac{L_2\,(6 - 9\,L_2 + 2\,{L_2}^2)}{6\,{L_1}^3} \\ &+ \frac{L_2\,(-12 + 36\,L_2 - 22\,{L_2}^2 + 3\,{L_2}^3)}{12\,{L_1}^4} + O\left(\left\{ \frac{L_2}{L_1} \right\}^5 \right) , \end{aligned} \tag{4.19}$$

where $L_1 = \operatorname{Log} z$ and $L_2 = \log \operatorname{Log} z$. This display of the expansion corrects a typographical error in equation (2.4.4) in [11].

To complete this development, it only remains to determine which branch of W is approximated when a particular branch of Log is chosen in (4.18). A straightforward analysis of the imaginary parts of the first two terms of the series (4.18), which is asymptotically $\arg z + 2\pi k$ for some k, yields the concise result

$$\begin{aligned} W_k(z) =& \log z + 2\pi i k - \log(\log z + 2\pi i k) \\ &+ \sum_{k=0}^{\infty} \sum_{m=1}^{\infty} c_{km} \log^m(\log z + 2\pi i k)(\log z + 2\pi i k)^{-k-m} . \end{aligned} \tag{4.20}$$

This analysis uses the fact that the branch index makes the approximation a discrete function of k; it is also a continuous function on any given branch, and therefore constant. Thus, if it holds asymptotically as $z \to \infty$, it holds everywhere in the region.

A similar but purely real-valued series is useful for the branch $W_{-1}(x)$ for small $x < 0$. We can get a real-valued asymptotic formula from the above by using $\log(-x)$ in place of $\mathrm{Log}(z)$ and $\log(-\log(-x))$ in place of $\log(\mathrm{Log}(z))$. A similar argument to that above arrives at precisely equation (4.13) for v, and establishes existence and convergence. Again, the coefficients of the expansion are exactly the same. This series is not useful for complex x as the branch cuts do not correspond.

Series expansions about the branch point

If we put $p = +\sqrt{2(ez+1)}$ in $We^W = z$, and expand in powers of $1+W$, we obtain

$$\frac{p^2}{2} - 1 = We^{1+W} = -1 + \sum_{k \geq 1} \left(\frac{1}{(k-1)!} - \frac{1}{k!} \right) (1+W)^k \ , \tag{4.21}$$

and then the series can be reverted to give

$$W(z) = \sum_{\ell=0}^{\infty} \mu_\ell p^\ell = -1 + p - \tfrac{1}{3}p^2 + \tfrac{11}{72}p^3 + \cdots \ . \tag{4.22}$$

This series converges for $|p| < \sqrt{2}$. It can be computed to any desired order from the following recurrence relations [17]:

$$\mu_k = \frac{k-1}{k+1} \left(\frac{\mu_{k-2}}{2} + \frac{\alpha_{k-2}}{4} \right) - \frac{\alpha_k}{2} - \frac{\mu_{k-1}}{k+1} \ , \tag{4.23}$$

$$\alpha_k = \sum_{j=2}^{k-1} \mu_j \mu_{k+1-j} \ , \qquad \alpha_0 = 2, \quad \alpha_1 = -1 \ , \tag{4.24}$$

where $\mu_0 = -1$ and $\mu_1 = 1$. This relation, which follows from

$$2pW = \left(\frac{p^2}{2} - 1 \right) \frac{d(1+W)^2}{dp} \ , \tag{4.25}$$

allows rapid calculation of the series, and provides a useful means of calculating W near the branch point.

For W_{-1} we take $p = -\sqrt{2(ez+1)}$, provided $\mathrm{Im}(z) \geq 0$. For $\mathrm{Im}(z) < 0$, this series with the negative square root gives instead good approximations for $W_1(z)$. This is what we meant earlier when we said that the expansion at the branch point $-1/e$ was complicated by the choice of locations for the branch cuts.

The relation between W_{-1} and W_0 near the branch point was investigated in [45] by Karamata, who studied the coefficients c_n in the power series

$$\mu = \sigma + \tfrac{2}{3}\sigma^2 + \tfrac{4}{9}\sigma^3 + \tfrac{44}{135}\sigma^4 + \cdots = \sum_{n \geq 1} c_n \sigma^n \ , \tag{4.26}$$

being the solution to

$$(1+\mu)e^{-\mu} = (1-\sigma)e^{\sigma}\,, \qquad \mu = \sigma + O(\sigma^2)\,. \tag{4.27}$$

This power series arises, for example, in the study of random graphs [41, page 323]. He tabulated c_n for $n \le 15$ and proved that

$$\frac{1}{n} \le c_n \le \frac{1}{n} + \frac{1-1/n}{n(1+1/2+\cdots+1/n)}\,. \tag{4.28}$$

In Maple V Release 3, the branch cuts for W are those described above. Prior to Maple V Release 2, only the real-valued inverse of the real function $\xi \mapsto \xi e^{\xi}$ on $[-1, \infty)$ was implemented, so the question of branches did not arise. In Maple V Release 2, the branch cuts were chosen to be $(-\infty, -1/e]$ and $(0, \infty)$, but the simplification of the asymptotic expansions resulting from moving the latter branch cut to the negative real axis was considered to be significant enough to change the branch cuts for W in Maple V Release 3.

5. Numerical Analysis

In [27], Euler made brief mention of the complex roots of $x = a^x$ when a is real, but the first person to explain how all values $W_k(x)$ could be calculated for real x was apparently Lémeray [51,52]. Then in [36] Hayes showed how to find all the values $W_k(x)$ when x is complex, and how to bound their real part. Wright made further studies, reported in [70], and then wrote a comprehensive paper [71] containing a detailed algorithm for the calculation of all branches.

The numerical analysis of the W function must take into account two contexts: fixed precision implementations, and *arbitrary* precision implementations, the latter being needed for computer algebra systems such as Maple. We first consider the conditioning of W, and then go on to examine methods for efficient approximation.

The standard theory of conditioning of function evaluation (see e.g. [16, p. 14]) gives the expression $C = xf'(x)/f(x)$ which estimates the relative effect on $y = f(x)$ of a small relative change in x. Here it is easy to see that for $y = W(z)$, the condition number C is

$$C = \frac{1}{1+W(z)}\,. \tag{5.1}$$

This number tells us the approximate relative effect of uncertainty in z. We see immediately that W is ill-conditioned near the branch point $z = -1/e$ (for W_0, W_{-1} and W_1), when $W(z) \approx -1$. Interestingly, there appears to be no ill-conditioning near the singularity $z = 0$ for any branch; however, this is a consequence of considering *relative* error.

We can in fact say more than the standard theory gives. If we have some approximation to $W(z)$, say $\widetilde{w}$, then we can define the *residual*

$$r = \widetilde{w}e^{\widetilde{w}} - z\,. \tag{5.2}$$

This residual is computable, though some extra precision may be necessary, particularly if $\widetilde{w}$ is large and near the imaginary axis. Note that since $z + r = \widetilde{w}e^{\widetilde{w}}$, then $\widetilde{w} = W(z+r)$ exactly. Further,

$$\widetilde{w} = W(z+r) = W(z) + \int_z^{z+r} W'(\zeta)\,d\zeta\,, \tag{5.3}$$

where the path of integration in the second equation does not cross any branch cut. We emphasize that we have exact equality here—our approximation $\widetilde{w}$ is the exact value of W for a slightly different argument, where the backward error is just r. Then, use of the fundamental theorem of calculus (or, in the real-variable case, the Mean Value Theorem) as above gives us an exact expression for the forward error, as well. If we can find a simple bound for $|W'(z)|$ we will then have a good forward error *bound*.

To find such a bound, we first look at the real case. It is obvious from (3.2), or from the graph of $W(x)$, that $W'(x) < 1$ if $x > 0$. Hence we can say immediately that $W(x) - \widetilde{w} = W'(x + \theta r)r$ (for some $0 \le \theta \le 1$) is bounded in magnitude by $|r|$, which provides a very convenient and computable error bound (it is very pessimistic, though, for very large z). This suggests examining the region in the z-plane where $W'(z)$ has magnitude 1. We thus put

$$W'(z) = \frac{1}{(1+W)e^W} = e^{i\psi} \tag{5.4}$$

and now try to find the curve or curves in the z-plane which satisfy that equation.

To do this, we first modify and then solve equation (5.4) exactly, as follows. We consider the more general case where we wish to find the curves where $W'(z)$ has magnitude ρ:

$$W'(z) = \frac{1}{(W(z)+1)e^{W(z)}} = \rho e^{i\psi} . \tag{5.5}$$

Invert, and then multiply both sides by e and notice that

$$1 + W(z) = W(\rho^{-1}e^{-i\psi}) \tag{5.6}$$

(for *some* branch of W) and use $z = We^W$ to get

$$z = \rho^{-1}e^{-i\psi}\left(1 - \frac{1}{W(\rho^{-1}e^{-i\psi})}\right) . \tag{5.7}$$

This gives us an exact and very interesting expression for the location of the curves where $|W'(z)| = \rho$. One can see immediately that the curves for $W_k(z)$ are the same as those for $W_{-k}(z)$, by using the near conjugate symmetry relation $\overline{W_k(z)} = W_{-k}(\overline{z})$ (see Section 4). Further, since the magnitude of $W_k(z)$ increases as k increases, we see that for large k the curve in the z-plane is really a small modification of the circle of radius ρ^{-1} centred around the branch point $z = 0$. Thus for large k one sees that essentially the forward error is less than 10 times the residual if z is outside a circle of radius $1/10$ centred around the origin.

This gives us a complete error analysis, because away from the branch points we can bound $W'(z)$ and thus compute a bound on the forward error from the computable backward error. However, the above analysis is not terribly practical if, for example, we are at all close to the branch point $z = -1/e$ and wish to compute $W_0(z)$, $W_{-1}(z)$ or $W_1(z)$. In these cases, though, we may use the series expansion (4.22) which is valid near the branch point itself, or Padé approximants based on that series.

The residual r of approximations based on n terms of (4.22) is $O(p^{n+1})$. Thus for z close enough to the branch point, the smallness of the residual itself compensates somewhat for the amplification of the forward error due to ill-conditioning near that point. For simplicity we abandon *error bounds* for *error estimates*, and make this compensation precise.

Equation (4.22) implies that near the branch point $W_0'(z) = (1-2/3p+O(p^2))dp/dz$. Now $p^2 = 2(ez+1)$, so $dp/dz = e/p$. This implies that $W_0'(z) = O(1/p)$ near the branch point. Thus if we are attempting to calculate $W_0(z_0)$ and $p_0 = \sqrt{2(ez_0+1)}$ is small, and further we have an estimate with residual $O(p_0^k)$ by taking $k-1$ terms in the series about the branch point, then the forward error is clearly $O(p_0^{k-1})$.

For definiteness, suppose we take $w = -1 + p_0 - \frac{1}{3}p^2$. Then the residual is given by $r = -11p_0^4/(72e) + O(p_0^5)$, and the forward error is

$$W_0(p_0) - W_0(p_0+r) = -\int_{p_0}^{p_0+r} W_0'(\zeta)\,d\zeta = \frac{11}{72}p_0^3 + O(p_0^4)\,. \tag{5.8}$$

Notice that this agrees, as it should, with simply comparing w with a higher-order accurate initial approximation. Thus we lose one order of accuracy of the series approximation in going from the residual to the forward error.

We turn now to practical methods for computing all the branches of $W(z)$. We are interested in such computations in an arbitrary precision context, and so we focus our attention on methods which are easily scaled to higher accuracy. Since W is an inverse function, it is natural to consider root-finding methods such as Newton's method, which is a second-order method. For general root-finding problems, there is little to be gained by considering higher order methods, because the cost of computing the requisite derivatives becomes prohibitive. However, for the function $x \mapsto xe^x$ this is not the case, because the nth derivative is just $(x+n)e^x$, which is obtainable at the cost of one (complex) floating point multiplication operation, once the value of e^x has been computed.

In a variable precision environment there is another very significant feature of iterative rootfinders: given an nth order method, once convergence begins, the number of digits correct at step k is roughly n times the number which were correct at step $(k-1)$, and this means that if the value of the root is to be computed to d digits of precision, then only the last iteration need be computed to d digits, while the penultimate iteration can be computed to d/n digits, the iteration before that to d/n^2 digits, and so on. Furthermore, if step $(k-1)$ is computed to d_{k-1} digits, then to obtain the kth estimate of the root, the correction term to be added need only be computed to $(n-1)d_{k-1}$ digits, since the first d_{k-1} digits will not be affected by this correction term, and the sum will then be correct to nd_{k-1} digits. This analysis is similar to that in [66], and we remark that the residual in the Newton-like method must be carefully computed to ensure the correction term is accurate.

To estimate the cost of such a scheme, suppose convergence begins with a 1 digit accurate initial guess. For an nth order scheme and for a root desired to d digits, approximately $\log_n d$ iterations will be required, with the precision increasing by a factor of n at each iteration, and much of the work at the kth iteration can be carried out at $(n-1)/n$ times the current working precision. There is thus a trade-off between the savings realized by the reduced precision required for intermediate computations, which are higher for lower order schemes (since $n/(n+1) > (n-1)/n$ for positive integer n), and the savings realized by the reduction in the total number of iterations required by higher order schemes. In general, the higher costs associated with the computation of higher order derivatives for higher order methods must also be considered.

We return to the specific problem at hand, that of computing a value of $W_k(z)$ for arbitrary integer k and complex z. Taking full advantage of the features of iterative rootfinders outlined above, we compared the efficiency of three methods, namely, (1) Newton's method, (2) Halley's method (see e.g. [1]), which is a third-order method, and (3) the fourth-order method described in [30] (as published, this last method evaluates only the principal branch of W at positive real arguments, but it easily extends to all branches and to all complex arguments). The methods were coded using Maple V Release 3, and the comparisons were done on a DEC (Alpha) 3000/800S. For the purposes of comparing methods, the initial guess algorithm is not significant, so long as the same one is used for all methods. The results showed quite consistently that method (2) is the optimal method in this environment, with method (3) generally second best and method (1) usually the slowest. These rankings were consistent across a wide range of precisions, branches and arguments, including principal branch evaluations at positive real arguments (for which the entire computation involves real arithmetic only). The rankings were also consistent with those obtained on other platforms.

It should be noted that in [30] Fritsch *et al.* chose to use a 4th order method because they were interested in obtaining an algorithm which could return 6 figure accuracy with just one application of the method, and consequently they focussed much of their attention on the initial guess algorithm. They also provided a hybrid 3rd/4th order method which could obtain machine accuracy on a CDC 6000 with two iterations. It would make sense to incorporate their work into a fast, low precision Maple evaluator for $W_0(x)$, $-1/e \le x$, and $W_{-1}(x)$, $-1/e \le x < 0$, but that has not yet been done. A project to develop a routine for the single and double precision evaluation of all of the branches of W at real and complex arguments is currently under way.

For the W function, Halley's method takes the form

$$w_{j+1} = w_j - \frac{w_j e^{w_j} - z}{e^{w_j}(w_j+1) - \dfrac{(w_j+2)(w_j e^{w_j} - z)}{2w_j+2}} \,. \tag{5.9}$$

In the Maple implementation, the initial guess is determined as follows. For most arguments z, a sufficiently accurate value is given by just the first two terms of the asymptotic expansion (4.18). When computing any of $W_0(z)$, $W_{-1}(z)$ or $W_1(z)$ for z near the branch point at $-1/e$, the first two terms of the series (4.22) can be used (with the appropriate changes of sign in the case of W_{-1} and W_1). When computing $W_0(z)$ near 0, a Padé approximant can be used (Maple uses a (3,2)-Padé approximant). Finally, when computing $W_{-1}(z)$ for z near but not too near either of the branch points 0 or $-1/e$ a simple rational approximation to W_{-1} should be used, as for such z neither the asymptotic expansion nor the series expansion at $-1/e$ is very accurate. A similar remark holds true for W_1. For the implementation in Maple, the approximate boundaries of these regions were determined empirically.

6. Concluding Remarks

We have collected here many available results on the Lambert W function, for convenient reference. We have presented some of the history of W and some examples of applications; we have presented new results on the numerical analysis, asymptotic analysis, and symbolic calculus of this function. An important part of this paper is our proposal of a standard notation for all the branches of W in the complex plane (and, likewise, for the related tree function $T(x)$). Names are important. The Lambert W function has been widely used in many fields, but because of differing notation and the absence of a standard name, awareness of the function was not as high as it should have been. Since the publication of this paper as a technical report and since the publication of [21], many more applications have been recognized.

Acknowledgements.

A very large number of people responded to the publication of this paper as a technical report, and we are very grateful for their contributions, which have substantially improved the work. We are especially indebted to André Heck and Neville Robinson for pointing out the connection to delay-differential equations. Bruno Salvy pointed out that the coefficients of the asymptotic series for W were known in closed form, and gave us the reference to Comtet's work. J. Borwein pointed out a connection between (3.11) and the Legendre Transform [4]. The jet-fuel problem was communicated to us by Michael Kamprath, with commentary by Harald Hanche-Olsen. We thank G.-C. Rota and D. Loeb for discussions, and we thank Prof. D. E. Gerber for his help in translating (from the Latin) parts of Euler's paper [28]; other parts of the paper were translated by Matt Davison. Finally, we thank Nathalie Sinclair, who translated Lambert's *éloge* from the French.

This work was supported in part by the Natural Science and Engineering Research Council of Canada and the Information Technology Research Council of Ontario.

References

[1] G. Alefeld, "On the convergence of Halley's method", Amer. Math. Monthly, **88** (1981) 530–536.

[2] American National Standards Institute/Institute of Electrical and Electronic Engineers: *IEEE Standard for Binary Floating-Point Arithmetic,* ANSI/IEEE Std 754-1985, New York, 1985.

[3] J.D. Anderson, *Introduction to Flight*, 3rd Ed., McGraw-Hill, New York, 1989.

[4] V.I. Arnold, *Mathematical Methods for Classical Mechanics*, Springer-Verlag, 1978.

[5] I.N. Baker and P.J. Rippon, "Convergence of infinite exponentials", Ann. Acad. Sci. Fenn. Ser. AI Math. **8** (1983) 179–186.

[6] I.N. Baker and P.J. Rippon, "A note on complex iteration", Amer. Math. Monthly, **92** (1985) 501–504.

[7] D.A. Barry, J.-Y. Parlange, G.C. Sander and M. Sivaplan, "A class of exact solutions for Richards' equation", J. Hydrology, **142** (1993) 29–46.

[8] R.E. Bellman and K.L. Cooke, *Differential-Difference Equations*, Academic Press, 1963.

[9] W.H. Beyer (ed.), *CRC Standard Mathematical Tables*, 28th Ed. 1987.

[10] C.W. Borchardt, "Ueber eine der Interpolation entsprechende Darstellung der Eliminations-Resultante", J. reine angewandte Math. **57** (1860) 111–121.
[11] N.G. de Bruijn, *Asymptotic Methods in Analysis*, North-Holland, 1961.
[12] C. Carathéodory, *Theory of Functions of a Complex Variable*, Chelsea, 1954.
[13] A. Cayley, "A theorem on trees", Quarterly Journal of Mathematics, Oxford Series, **23** (1889) 376–378.
[14] B.W. Char, K.O. Geddes, G.H. Gonnet, B.L. Leong, M.B. Monagan and S.M. Watt, *The Maple V Language Reference Manual*, Springer-Verlag, 1991.
[15] L. Comtet, *Advanced Combinatorics*, D.Reidel, 1974.
[16] S.D. Conte and C. de Boor, *Elementary Numerical Analysis*, 3rd Ed., McGraw-Hill, 1980.
[17] D. Coppersmith, private communication.
[18] R.M. Corless, "What good are numerical simulations of chaotic dynamical systems?", Computers Math. Applic. **28** (1994) 107-121.
[19] R.M. Corless, *Essential Maple*, Springer-Verlag 1994.
[20] R.M. Corless and D.J. Jeffrey, "Well, It Isn't Quite That Simple", SIGSAM Bulletin, **26**, 3 (1992) 2–6.
[21] R.M. Corless, G.H. Gonnet, D.E.G. Hare, and D.J. Jeffrey, "Lambert's W function in Maple", The Maple Technical Newsletter, **9** (1993) 12–22.
[22] H.T. Davis, *Introduction to Nonlinear Differential and Integral Equations*, Dover, 1962.
[23] L. Devroye, "A note on the height of binary search trees", Journal of the ACM, **33** (1986) 489 –498.
[24] O. Dziobek,"Eine Formel der Substitutionstheorie", Sitzungsberichte der Berliner Mathematischen Gesellschaft, **17** (1917), 64–67.
[25] G. Eisenstein, "Entwicklung von $\alpha^{\alpha^{\cdot^{\cdot^{\cdot}}}}$", J. reine angewandte Math. **28** (1844) 49–52.
[26] P. Erdős and A. Rényi, "On the evolution of random graphs", Magyar Tud. Akad. Mat. Kut. Int. Közl. **5** (1960) 17–61. Reprinted in P. Erdős, *The Art of Counting* (1973), pp 574–618, and in *Selected Papers of Alfréd Rényi* (1976), pp 482–525.
[27] L. Euler,"De formulis exponentialibus replicatis", Leonhardi Euleri Opera Omnia, Ser. 1, Opera Mathematica **15** (1927) [original date 1777], 268–297.
[28] L. Euler, "De serie Lambertina plurimisque eius insignibus proprietatibus", Leonhardi Euleri Opera Omnia, Ser. 1, Opera Mathematica, **6**, 1921 (orig. date 1779), 350-369.
[29] P. Flajolet and M. Soria, "Gaussian Limiting Distributions for the Number of Components in Combinatorial Structures", J. Combinatorial Theory, Series A, **53** (1990) 165–182.
[30] F.N. Fritsch, R.E. Shafer, and W.P. Crowley, "Algorithm 443: Solution of the transcendental equation $we^w = x$", Comm. ACM, **16** (1973) 123–124.
[31] K.O. Geddes, S.R. Czapor and G. Labahn, *Algorithms for Computer Algebra*, Kluwer Academic Publishers, 1992.
[32] Ch.C. Gillispie, ed., *Dictionary of Scientific Biography*, Scribners, N. Y., 1973.
[33] G.H. Gonnet, *Handbook of Algorithms and Data Structures* Addison-Wesley, 1984.
[34] G.H. Gonnet, "Expected length of the longest probe sequence in hash code searching", Journal ACM, **28** (1981) 289 – 304.

[35] R.L. Graham, D.E. Knuth, O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 1994.

[36] N.D. Hayes, "Roots of the transcendental equation associated with a certain difference-differential equation", J. Lond. Math. Soc. **25** (1950) 226 – 232.

[37] N.D. Hayes, "The roots of the equation $x = (c\exp)^n x$ and the cycles of the substitution $(x|ce^x)$", Q. J. Math. **2** (1952) 81–90.

[38] T.L. Heath, *A Manual of Greek Mathematics*, Dover, 1963.

[39] E.W. Hobson, *Squaring the Circle*, Chelsea Publishing Company, 1953.

[40] T.E. Hull, W.H. Enright, B.M. Fellen, and A.E. Sedgwick, "Comparing numerical methods for ordinary differential equations", SIAM J. Numer. Anal. **9** (1972) 603–637.

[41] S. Janson, D.E. Knuth, T. Łuczak, and B. Pittel, "The birth of the giant component", *Random Structures and Algorithms*, **4** (1993) 233–358.

[42] D.J. Jeffrey, R.M. Corless, D.E.G. Hare & D.E. Knuth, "Sur l'inversion de $y^\alpha e^y$ au moyen de nombres de Stirling associés", C. R. Acad. Sc. Paris, Série I, **320**, 1449–1452.

[43] D.J. Jeffrey, R.M. Corless, D.E.G. Hare, "Unwinding the branches of the Lambert W function", The Mathematical Scientist, *in press*

[44] W. Kahan,"Branch cuts for complex elementary functions", In *The State of the Art in Numerical Analysis:* Proceedings of the Joint IMA/SIAM Conference on the State of the Art in Numerical Analysis, University of Birmingham, April 14-18, 1986, M. J. D. Powell and A. Iserles, Eds, Oxford University Press.

[45] J. Karamata, "Sur quelques problèmes posés par Ramanujan", J. Indian Math. Soc. **24** (1960) 343–365.

[46] R.A. Knoebel, "Exponentials reiterated", Amer. Math. Monthly, **88** (1981) 235–252.

[47] D.E. Knuth and B. Pittel, "A recurrence related to trees", Proc. Amer. Math. Soc. **105** (1989) 335–349.

[48] J.H. Lambert, "Observationes variae in mathesin puram" Acta Helvetica, physico-mathematico-anatomico-botanico-medica, **3**, Basel, (1758) 128–168.

[49] J.H. Lambert, "Observations Analytiques," in *Nouveaux mémoires de l'Académie royale des sciences et belles-lettres,* Berlin, 1772, vol. 1, for 1770.

[50] H.G. Landau, "On some problems of random nets", Bull. Mathematical Biophysics, **14** (1952) 203 – 212.

[51] E.M. Lémeray, "Sur les racines de l'equation $x = a^x$", Nouvelles Annales de Mathématiques (3) **15** (1896) 548–556.

[52] E.M. Lémeray, "Sur les racines de l'equation $x = a^x$. Racines imaginaires", Nouvelles Annales de Mathématiques (3) **16** (1897) 54–61.

[53] E.M. Lémeray, "Racines de quelques équations transcendantes. Intégration d'une équation aux différences mèlées. Racines imaginaires", Nouvelles Annales de Mathématiques (3) **16** (1897) 540–546.

[54] R.E. O'Malley, Jr., *Singular Perturbation Methods for Ordinary Differential Equations*, Springer-Verlag Applied Mathematical Sciences, **89**, 1991.

[55] F.D. Parker, "Integrals of inverse functions", Amer. Math. Monthly, **62**, (1955) 439–440.

[56] G. Pólya and G. Szegö, *Problems and Theorems in Analysis*, Springer-Verlag, 1972.

[57] G. Pólya, "Kombinatorische Anzahlbestimmungen für Gruppen, Graphen und chemische Verbindungen", Acta Mathematica **68** (1937) 145–254. English translation by Dorothee Aeppli in G. Pólya and R.C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds*, Springer-Verlag, 1987.

[58] K.B. Ranger, "A complex variable integration technique for the 2-dimensional Navier–Stokes equations", Q. Applied Maths, **49** (1991) 555–562.

[59] E.L. Reiss, "A new asymptotic method for jump phenomena", SIAM J. Appl. Math. **39** (1980) 440–455.

[60] J.M. Robson, "The height of binary search trees", Aust. Comput. J. **11** (1979) 151–153.

[61] L.A. Segel and M. Slemrod, "The quasi-steady-state assumption: a case study in perturbation", SIAM Review, **31** (1989) 446–477.

[62] T.C. Scott, J.F. Babb, A. Dalgarno and J.D. Morgan III, "Resolution of a paradox in the calculation of exchange forces for H_2^+", Chem. Phys. Lett. **203** (1993) 175–183.

[63] T.C. Scott, J.F. Babb, A. Dalgarno and J.D. Morgan III, "The calculation of exchange forces: general results and specific models", J. Chem. Phys. **99** (1993) 2841– 2854.

[64] O. Skovgaard, I.G. Jonsson, and J.A. Bertelsen, "Computation of wave heights due to refraction and friction", J. Waterways Harbours and Coastal Engineering Division, February, 1975, pp 15–32.

[65] R. Solomonoff and A. Rapoport, "Connectivity of random nets", Bull. Math. Biophysics, **13** (1951) 107–117.

[66] D.C. Sorensen and Ping Tak Peter Tang, "On the orthogonality of eigenvectors computed by divide-and-conquer techniques", SIAM J. Numer. Anal. **28** no. 6, (December 1991) 1752–1775.

[67] J.J. Sylvester, "On the change of systems of independent variables", Q. J. Pure and Applied Maths, **1** (1857) 42–56.

[68] E.C. Titchmarsh, *Theory of Functions*, 2nd ed., Oxford, 1939.

[69] E.M. Wright, "The linear difference-differential equation with constant coefficients," Proc. Royal Soc. Edinburgh, A**62** (1949) 387–393.

[70] E.M. Wright, "A non-linear difference-differential equation", J. für reine und angewandte Mathematik, **194** (1955) 66 – 87.

[71] E.M. Wright,"Solution of the equation $ze^z = a$", Proc. Roy. Soc. Edinburgh, (A)**65** (1959) 193–203.

[72] E.M. Wright,"The number of connected sparsely edged graphs", J. Graph Theory, **1** (1977) 317–330.

Appendix: Biographical Notes.

Johann Heinrich Lambert was born in Mulhouse on 26 August 1728, and died in Berlin on 25 September 1777 [32]. His scientific interests were remarkably broad. The self-educated son of a tailor, he produced fundamentally important work in number theory, geometry, statistics, astronomy, meteorology, hygrometry, pyrometry, optics, cosmology and philosophy. He worked on the parallel postulate, and also introduced the modern notation for the hyperbolic functions. It is said that when Frederick the Great asked him in which science he was most proficient, Lambert immodestly replied, "All." Lambert was the first to prove the irrationality of π. We find it a remarkable coincidence that the curves defining the branch cuts of the Lambert W function (which contain the Quadratix of Hippias) can be used not only to square the circle—which, by proving π irrational, Lambert went a long way towards proving was impossible by compass and straightedge—but also to trisect a given angle [39, 38].

The title of Euler's paper [28] can be translated as *On a series of Lambert and some of its significant properties*. Indeed, Euler refers to Lambert in this paper as "acutissimi ingenii Lambertus", which may be freely translated as *the ingenious engineer Lambert*. Although Euler was Swiss, he evidently did not read Acta Helvetica regularly, because Lambert's formulae came as a big surprise when he learned of them in 1764, which was the year Lambert went from Zürich to Berlin. A letter from Euler to Goldbach, dated 17 March 1764, describes his excitement.

In [49], Lambert states for the first time the generalization of his series that gives powers of the root instead of just the root itself (this series is equation (1.2) here). Lambert says that *Acta Helvetica* had cut out the proof of his simpler formula, and he had lost his notes. But while rederiving the formula for Euler he found a simpler proof of a more general result. Then he says Euler worked out a generalization from trinomials to quadrinomials; but he (Lambert) already knew how to handle polynomials, and to derive a precursor of Lagrange's inversion theorem. Lambert wrote Euler a cordial letter on 18 October 1771, hoping that Euler would regain his sight after an operation; he explains in this letter how his trinomial method extends to series reversion. In the index to Euler's correspondence [Opera Omnia, series 4, volume 1, page 246] there is a note that A. J. Lexell replied to this letter.

Finally, in an *éloge* near the beginning of the two volumes of Lambert's mathematical papers (published in Zürich in 1946 and 1948) we find that Lambert had dreams of building a machine to do symbolic calculation (while Pascal's machine merely did arithmetic). Thus his wishes anticipated those of Lady Ada Lovelace by roughly one century, and the actuality of symbolic computation systems by roughly two.