Editor's Corner: The Unwinding Number

Robert M. Corless and David J. Jeffrey

SIGSAM Bulletin 116, 26 July 1996, pp 28 – 35

1 Introduction

From the Oxford English Dictionary we find that *to unwind* can mean "to become free from a convoluted state". Further down we find the quotation "The solution of all knots, and unwinding of all intricacies", from H. Brooke (The Fool of Quality, 1809). While we do not promise that the unwinding number, defined below, will solve *all* intricacies, we do show that it may help for quite a few problems.

Our original interest in this area came from a problem in which an early version of DERIVE was computing the wrong answer when simplifying $\sin(\sin^{-1} z)$, which should always be just z. For $z > 1$, DERIVE was getting $-z$ as the answer. This bug has of course long since been fixed.

What was happening was that in order to improve internal efficiency, all the inverse trig functions were represented as arctangents. Consulting an elementary book of tables, one finds the identity

$$\sin^{-1} z = \tan^{-1}\left(\frac{z}{\sqrt{1-z^2}}\right) . \tag{1}$$

In the same vein, one finds that

$$\sin(\tan^{-1} w) = \frac{w}{\sqrt{1+w^2}} . \tag{2}$$

Substituting equations (1) and (2) into $\sin(\sin^{-1} z)$ and simplifying, we get

$$\frac{z}{\sqrt{1-z^2}} \frac{1}{\sqrt{\dfrac{1}{1-z^2}}} , \tag{3}$$

which DERIVE quite properly refused to simplify to z, because this is *not* always equal to z (see [2]).

The fix in this case was to replace equation (2) with

$$\sin(\tan^{-1} w) = w\sqrt{\frac{1}{1+w^2}} . \tag{4}$$

which differs from the original only on the branch cut. See [7] for more discussion. This change allows the simplification of $\sin(\sin^{-1} z)$ to z. Verifying that this approach worked, and indeed trying to understand the problem to begin with, led us to attempt various definitions of a 'branch function'. This introductory problem turned out to be the tip of an iceberg of problems connected with using the principal branch of multivalued elementary functions.

1.1 Logs and Branches

In what follows, the *principal branch of the logarithm function* is denoted by $\ln z = \ln|z| + i\theta$ with $-\pi < \theta = \arg z \leq \pi$, the now-conventional closure on the top of the branch cut (known as Counter-Clockwise Continuity or CCC [6]). Any other branch choice would lead to a similar discussion. In section 2 we discuss the option of not choosing branches at all by using Riemann surfaces.

The very idea of choosing the principal branch of the logarithm, or indeed choosing a consistent single-valued branch of the logarithm at all, has some unpleasant consequences for computer algebra, in that we lose several algebraic identities that we would like to use automatically. These consequences are unpleasant for humans as well, because it is hard to unlearn algebraic rules that we use so nearly automatically ourselves.

Some identities remain true, of course, and here are some representatives.

1. $\exp(\ln z) = z$
2. $\exp(z_1 + z_2) = \exp(z_1)\exp(z_2)$
3. $\exp(2\pi i k) = 1$.

The algebraic rules we lose include

1. $\ln(z_1 z_2) \neq \ln(z_1) + \ln(z_2)$
2. $(z_1 z_2)^\alpha \neq z_1^\alpha z_2^\alpha$
3. $\ln(z^w) \neq w \ln z$
4. $(z^\alpha)^\beta \neq z^{\alpha\beta}$, and in particular $(z^\alpha)^\beta \neq (z^\beta)^\alpha$, which takes quite a bit of getting used to. We adopt the convention that $z^{\alpha\beta}$ means $\exp(\alpha\beta \ln z)$, while the parenthesized symbols have the meanings implied by the precedence of the operations.

The purpose of this column is to see exactly how these identities have to be modified, once we choose the principal branch of the logarithm. Introducing the unwinding number $\mathcal{K}(z)$ turns out to be sufficient for this purpose.

1.2 Unwinding number

We define the *unwinding number* $\mathcal{K}(z)$ by

$$\ln(e^z) = z + 2\pi i\mathcal{K}(z)\ . \tag{5}$$

See [4], where this function is used to derive new identities for the Lambert W function.

Functions similar to $\mathcal{K}$ have been defined several times in the literature. In 1974, Apostol [1] briefly considered a cognate of $\mathcal{K}$. Charles Patton has defined several functions including $\mathrm{UNLN}(z) = \ln\exp z - z$ (see [8] for a brief discussion of UNLN) which is $2\pi i\mathcal{K}(z)$ in our notation. Aslaksen [2] defines several functions including $\mathrm{Imq}(z)$, which turns out to be $-\mathcal{K}(z)$ in our notation. It would be interesting to see the results of a thorough historical investigation.

One can define $\mathcal{K}(z)$ without logarithms by using the floor function. If $\Im(z)$ is the imaginary part of z, then

$$\mathcal{K}(z) = \mathcal{K}(i\Im(z)) = \left\lfloor \frac{\pi - \Im(z)}{2\pi} \right\rfloor\ . \tag{6}$$

It is easy to see that $\mathcal{K}(z) = 0$ if $-\pi < \Im(z) \le \pi$, and in general that $\mathcal{K}(z) = -n$ if $(2n-1)\pi < \Im(z) \le (2n+1)\pi$. Thus the unwinding number is constant on horizontal strips. Note the closure on the top of the strips.

The function was called the 'unwinding number' because we thought of $\exp z$ as winding z around the branch point of log; in order to get z back one has to 'unwind'.

2 Connection with the Riemann surface for logarithm

Is it necessary introduce a new function at all? Surely the properties of the logarithm function are well understood by now? This is of course perfectly true, but some apparently minor things have changed since the theory of the complex-valued logarithm function was first elucidated. These are

1. the rise of computers and the concomitant increased need for the single-valued (numerical) complex logarithm,

2. the establishment of a consensus (articulated for example in [6]) on where to close the branch cut for the principal branch of the logarithm ($-\pi < \arg z \le \pi$), and

3. the creation of symbolic manipulation languages that manipulate formulas algebraically, leaving numerical evaluation as late as possible.

Choosing a branch of logarithm may introduce instances of the so-called *specialization problem*, wherein a formula that is right most of the time can be wrong for special values of the input. Introducing $\mathcal{K}$ fixes this.

But perhaps we should not invent a new function if there is an existing theory designed to deal with the multivalued nature of the logarithm, which is the ultimate source of the difficulty here.

Let us consider the possibilities offered by a *Riemann surface*. Consult practically any complex analysis textbook for a discussion of this idea. Basically, we deal with the multiple covering of $\mathbb{C}$ by $\exp z$ by considering a new set $\mathcal{R}$ which is to be the range for the exponential function. We denote this slightly different exponential function by $\exp_{\mathcal{R}} z$, as its range is different and would require a different data structure in an implementation. Classically $\mathcal{R}$ is a helix consisting of a countable infinity of copies of the complex plane, each cut along the negative real axis and joined to the sheets immediately above and immediately below. Once the joins are made 'invisible' one can show that the function $\exp_{\mathcal{R}} z$ is one-to-one and analytic on this surface, and thus has a unique analytic inverse. We will denote this inverse function by $\log_{\mathcal{R}}(z)$ to distinguish it from the principal branch logarithm $\ln z$.

One can use polar coordinates (r,θ) on $\mathcal{R}$, where now we do not take θ modulo 2π. This provides a natural way of defining multiplication, as $z_1z_2 = (r_1r_2, \theta_1+\theta_2)$. If $p = \exp_{\mathcal{R}} z$ takes values on $\mathcal{R}$, then $p = (\exp(x), y)$ if $z = x+iy$, whilst $\log_{\mathcal{R}}(r,\theta) = \ln r + i\theta$. Representing (r,θ) in Cartesian coordinates requires *three* items, (x,y,k) where $x = r\cos\theta$, $y = r\sin\theta$, and $\theta = 2\pi k + \theta_0$ with $-\pi < \theta_0 \le \pi$ chosen for compatibility with the conventional principal branch cut. The integer k can be thought of as the index of the Riemann sheet on which (r,θ) lies. Thus we see that computation with elements of a Riemann surface still seems to require a choice of representation of the fundamental angle.

Note that there is a relation between $\log_{\mathcal{R}}$ and ln. If $p = x+iy$ and (p,k) denotes the Cartesian representation of (r,θ) then with the obvious meanings

$$\log_{\mathcal{R}}(p,k) = \ln p + 2\pi i k\ . \tag{7}$$

Suppose now $(p,k) = \exp_{\mathcal{R}}(z_0 + 2\pi i k)$, so $p = \exp(z_0)$, with $-\pi < \Im(z_0) \le \pi$. Then

$$\begin{aligned}\log_{\mathcal{R}}(p,k) &= z = z_0 + 2\pi i k \\ &= \ln(\exp z_0) + 2\pi i k\ .\end{aligned}$$

Rearranging this we have, since $\exp z = \exp z_0$,

$$\ln\exp z = z - 2\pi i k = z + 2\pi i\mathcal{K}(z)\ , \tag{8}$$

where k is the index of the Riemann sheet. That is, we may *interpret* the unwinding number $\mathcal{K}(z)$ as the negative of the index of the Riemann sheet.

Remark We are not really proposing this here, but we would like to see an implementation of functions of a Riemann-surface variable, to see if it offers any advantages, once the initial effort of constructing the various function representations has been made. There are increased costs for arithmetic on Riemann surfaces, and it is not clear what to do with addition, for example, or iterated functions (we would like $\ln\ln\exp\exp z$ to be z, for example, but this seems to require a "second order" Riemann surface). However, note that all the usual identities of logarithms and powers are preserved if we work on this Riemann surface: $\log_{\mathcal{R}}(\exp z) = z$, $(z_1 z_2)^\alpha = z_1^\alpha z_2^\alpha$, etcetera. This would make the symbolic algebra very simple indeed.

While this is interesting, and might be practicable, we think that learning how to do symbolic algebra in the existing domain with a principal-branch logarithm is still worthwhile, and it is for this reason that we have introduced $\mathcal{K}(z)$, which as we have seen has some connection with the theory of Riemann surfaces anyway.

3 The clearcut region

Given a function $f(z)$, the region where $\mathcal{K}(f(z)) = 0$ is often of particular importance. Define

$$\operatorname{clearcut}(f) := \{z \mid \mathcal{K}(f(z)) = 0\} \tag{9}$$

as the clearcut region[1] of f. The association of 'clearcut' with logs may make this somewhat mnemonic, and we will find that in this region the algebra is dramatically simpler (i.e. more 'clear cut') than otherwise.

The region is very simple to compute, given a function f. One simply finds the values of z that give $-\pi < \Im(f(z)) \le \pi$, if any, and this is the clearcut region for f. For example, if $f(z) = \ln z$, then $\operatorname{clearcut}(f)$ is the set where $\Im(\ln z) = \arg z$ is in $-\pi < \arg z \le \pi$. But this is in fact the entire complex plane. We have thus shown that for all z,

$$\mathcal{K}(\ln z) = 0\,. \tag{10}$$

We will use this process on several elementary functions in section (6).

Note that if f is real for $x \in D \subset \mathcal{R}$ real then $D \subset \operatorname{clearcut}(f)$.

4 Useful Theorems

Introducing a new function is all very well, but we need to be able to *do* things with it. The following theorems provide some algebraic rules for the manipulation of $\mathcal{K}$.

1. **Theorem**

$$\mathcal{K}(z + 2\pi i n) = \mathcal{K}(z) - n$$

for integer n. The proof is obvious from the definition or graph of $\mathcal{K}(z)$.

2. **Theorem**

$$\mathcal{K}(\ln z) = 0\,.$$

We proved this in the example in section (3).

3. **Theorem**

$$\ln(z_1 z_2) = \ln z_1 + \ln z_2 + 2\pi i \mathcal{K}\left(\ln z_1 + \ln z_2\right)\,.$$

To prove this, start with $z_1 z_2 = \exp(\ln z_1 + \ln z_2)$, and take logarithms to get

$$\begin{aligned} \ln(z_1 z_2) &= \ln(e^{\ln z_1 + \ln z_2}) \\ &= \ln z_1 + \ln z_2 \\ &\quad + 2\pi i \mathcal{K}(\ln z_1 + \ln z_2) \end{aligned}$$

by the definition of $\mathcal{K}$.

4. **Theorem** (generalization of Theorem 3)

$$\begin{aligned} \ln \prod_{k=1}^{n} z_k &= \sum_{k=1}^{n} \ln z_k \\ &\quad + 2\pi i \mathcal{K}\left(\sum_{k=1}^{n} \ln z_k\right)\,. \end{aligned}$$

To prove this we use induction. The case $n = 1$ is just Theorem 2, whilst the case $n = 2$ is just Theorem 3. Assuming the truth of the theorem for $n = m$, we have by Theorem 3 that

$$\begin{aligned} \ln(z_{m+1} \prod_{k=1}^{m} z_k) &= \ln z_{m+1} + \ln \prod_{k=1}^{m} z_k \\ &\quad + 2\pi i \mathcal{K}\left(\ln z_{m+1} + \ln \prod_{k=1}^{m} z_k\right) \end{aligned}$$

and using the inductive assumption to write

$$\begin{aligned} \ln \prod_{k=1}^{m} z_k &= \sum_{k=1}^{m} \ln z_k \\ &\quad + 2\pi i \mathcal{K}\left(\sum_{k=1}^{m} \ln z_k\right)\,, \end{aligned}$$

both inside the unwinding number and out, and further using Theorem 1 to cancel the inner unwinding number with the outer, we get the desired result.

[1] Thanks to Sumaya Corless for this name.

5. **Theorem** These all follow on writing a^b as its definition $\exp(b\ln(a))$:
   (a) $\ln(z^w) = w\ln z + 2\pi i\mathcal{K}(w\ln z)$
   (b) $(z_1z_2)^w = z_1^w z_2^w \exp(2\pi i w\mathcal{K}(\ln z_1 + \ln z_2))$ (the generalization to n terms in the product is immediate)
   (c) $(z^v)^w = z^{vw}\exp(2\pi i w\mathcal{K}(v\ln z))$ (notice that the order is important, and we ascribe our conventional meaning to z^{vw}).

5 Applications

In this section we give some sample applications, to show that this is not just an empty definition.

5.1 Fateman's z^w problem

Consider

$$y = z^w \tag{11}$$

as an equation for z, given y and w in $\mathbb{C}$, as discussed in [5]. We divide this into two problems: we first try to decide when $\zeta = y^{1/w}$ solves equation (11). This will give sufficient conditions for the classical formula to be true. We then try to discover all roots of (11), which turns out to be harder.

5.1.1 Sufficient conditions

Let $\zeta = y^{1/w}$. Then $\zeta^w = \exp(w\ln\zeta)$ or

$$\begin{aligned}\zeta^w &= \exp(w\ln\exp(\frac{1}{w}\ln y))\\ &= \exp(w(\frac{1}{w}\ln y + 2\pi i\mathcal{K}(\frac{1}{w}\ln y))\\ &= y\exp(2\pi i w\mathcal{K}(\frac{1}{w}\ln y))\ .\end{aligned}$$

This is equal to y if and only if $w\mathcal{K}((\ln y)/w)$ is an integer, say n. If w (which is given for the problem) is irrational, then n and hence $\mathcal{K}$ must be zero. If w is rational, then one can show by pigeonhole arguments that $\mathcal{K}$ must still be zero.

So ζ is a root of $z^w = y$ if and only if $\mathcal{K}((\ln y)/w) = 0$, or y is in the clearcut region for $(\ln y)/w$. This can happen if and only if $(\ln y)/w = t + ip$ where $-\pi < p \le \pi$. This implies that $\ln y = w(t+ip) = (a+ib)(t+ip) = (at-bp) + i(bt+ap)$ or, with $\ln y = \ln s + i\theta$ giving the polar coordinates of y, $s = \exp(at-bp)$ and $\theta = bt + ap$. If $b \ne 0$ we can eliminate the parameter t to get

$$s = e^{a\theta/b-(a^2+b^2)p/b}\ .$$

Remembering that $-\pi < p \le \pi$, then, if $ab \ne 0$, this is a domain bounded by logarithmic spirals. For Fateman's example $a = b = 1$ we have $s = \exp(\theta - 2p)$, which agrees with his plots (remember also that $-\pi < \theta \le \pi$).

Thus the obvious formula $\zeta = y^{1/w}$ for a root of $y = z^w$ is valid if and only if $\mathcal{K}((\ln y)/w) = 0$, which is a spiral (if $ab = \Im(w) \ne 0$; it is a sector if $b = 0$ and an annulus if $a = 0$) in the complex y-plane.

This treatment is shorter than Fateman's first-principles treatment, but remember that we have spent some time with a new function $\mathcal{K}(z)$. This time ought to pay off in practice, or else it isn't worthwhile.

5.1.2 Necessary conditions

Starting from equation (11), and assuming that a solution exists, we take logs. We find

$$\begin{aligned}\ln(y) &= \ln(z^w) = \ln(\exp(w\ln z))\\ &= w\ln z + 2\pi i\mathcal{K}(w\ln z)\end{aligned}$$

and so, rearranging, we find

$$\ln z = \frac{1}{w}(\ln(y) - 2\pi i\mathcal{K}(w\ln z)) \tag{12}$$

which is an implicit equation for z. Exponentiating both sides we get

$$z = y^{1/w}e^{-2\pi i\mathcal{K}(w\ln z)/w}\ . \tag{13}$$

We have thus shown that any solution must be of the form

$$z = y^{1/w}e^{-2\pi ik/w} \tag{14}$$

for some integer k. It will be a solution if and only if $\mathcal{K}(w\ln z) = k$. We can, with some more work, show using $\mathcal{K}$ that no solutions exist for some equations of this type, for example $z^{2/13} = i$. (The solution on the Riemann surface is $z = (1, 13\pi/4)$, but there is no ordinary complex z which solves this equation). We can show that for still other examples more than one solution may exist, even infinitely many (as in $z^{-i} = i$, for example), and $\mathcal{K}$ can help us with the algebra in this case also.

One can show in particular that if

$$\mathcal{K}((\ln y)/w - 2\pi ik/w) = 0 \tag{15}$$

then z from equation (14) solves equation (11), and that if w is irrational then this condition is also necessary. If $w = p/q$ is rational then things get more complicated.

However, it is not clear to us that when a user asks for the solution of $z^w = y$ that she really means 'find all the values of z for which this equation is true where z^w is defined to be $\exp(w\ln z)$'; that is, a user who realizes the difference between this and z^w on a Riemann surface is very sophisticated indeed. Finally, the formula $y^{1/w}$ on the Riemann surface is the *unique* answer to the problem there; in some sense what might make this formula fail here is the verification step. That is, $y^{1/w}$ may very well be the 'correct answer' even though $(y^{1/w})^w$ may not be equal to y, i.e. it doesn't satisfy the original equation when we use the principal branch to take the powers.

So for us the use of the unwinding number in this example is really just finding the clearcut region where we can use the classical formula and at the same time verify that the answer is correct by substituting it back in to the original equation.

5.2 The Aslaksen tests

Let us consider now how a computer algebra system that knew about the unwinding number might do on the Aslaksen tests [2]. We will presume the following model.

1. logarithms are expanded using $\mathcal{K}$.
2. powers are defined with ln and expanded using $\mathcal{K}$.
3. unwinding numbers are simplified using signum, csgn, and assumed information, and left alone if nothing can be inferred.

We now consider evaluation of expressions defined in some of the Aslaksen tests. We will consider in each case the consequences of natural assumptions on each variable.

1. $\sqrt{zw}$. By Theorem (5c) we would expect this to expand to
$$\sqrt{z}\sqrt{w}e^{\pi i \mathcal{K}(\ln z + \ln w)}$$
and this would not simplify further unless the assume system knew that $-\pi < \arg z + \arg(w) \leq \pi$, in which case $\mathcal{K}$ would simplify to 0. This would happen if both z and w were assumed positive, for instance.

2. $\sqrt{z^2}$. This case is similar to above, except the unwinding number that arises is $\mathcal{K}(2\ln z)$ and if $\arg z$ was known to be in the interval $(-\pi/2, \pi/2]$ then the unwinding number could be replaced with zero.

3. $\sqrt{1/z} - 1/\sqrt{z}$. We would expect that $\sqrt{1/z}$ would simplify according to the same rule as above, giving $\exp(\pi i \mathcal{K}(-\ln z))/\sqrt{z}$. The clearcut region for $-\ln z$ is the entire cut plane, except the negative real axis. So if z was known not to be real and negative, this would simplify to zero.

4. $\sqrt{e^z} - e^{z/2}$. Using our rules again, $\sqrt{\exp z} = \exp(\ln(\exp z)/2) = \exp(z/2 + \pi i \mathcal{K}(z))$. This would simplify to $\exp(z/2)$ if for example z was known to be real.

We leave the remaining tests as food for the reader's thought.

5.3 $\mathcal{K}$ and the Lambert W function

The unwinding number is not just useful for explaining how to modify identities so they work over the complex plane, but it has also been used to prove a new identity for the Lambert W function. This identity is that $W_k(z) + \ln W_k(z)$ is equal to

$$\begin{cases} \ln z & \text{if } k = -1 \text{ and } -1/e \leq z < 0\,, \\ \ln z + 2\pi i k & \text{otherwise}\,. \end{cases} \tag{16}$$

For the proof, see [4]. For a review article about W, see [3]. Briefly, $W_k(z)$ is the kth branch of the function satisfying $W(z)\exp(W(z)) = z$.

Incidentally, the clearcut region for $W_0(z)$ is the entire complex plane, because the range of $W_0(z)$ is wholly contained in the strip $-\pi < \Im(W_0(z)) < \pi$. The only other branches that have any nontrivial clearcut region are $W_{\pm 1}(z)$, but for $|k| > 1$ there is no z such that $\mathcal{K}(W_k(z)) = 0$. This means that it is never true that $\ln\exp W_k(z) = W_k(z)$ for $|k| > 1$.

6 Bestiary

In this section we present graphs of the clearcut regions for some of the simplest elementary functions. The graphs may be useful in and of themselves (for example we will learn that we may nearly always replace $\ln\exp\tan^{-1} z$ by $\tan^{-1} z$, but for z very near to the branch points at $\pm i$ this is incorrect), but the main intention of this section is to give examples of how to find the clearcut region for the problem you run into. We hope that these hand procedures will be formalized and implemented in a computer algebra system, of course. We have already seen the clearcut regions for z, $\ln z$, $-\ln z$, and $w\ln z$.

6.1 clearcut(z^n)

When is $\ln\exp(z^n) = z^n$, n an integer? This requires z to be in the clearcut region for z^n. Now $\mathcal{K}(z^n) = 0$ precisely when $-\pi < \Im(z^n) \leq \pi$. When $n = 2$ for example, $\Im(z^2) = 2xy$ and this becomes $-\pi/2 < xy \leq \pi/2$. These are right hyperbolae containing both the real and imaginary axes. When $n = 3$ we have $-\pi < 3x^2y - y^3 \leq \pi$ and this is a 6-pointed starlike region with thinning branches going off to infinity along the real axis and the lines $y = \pm\sqrt{3}x$. See Figure 1. Investigating a few more convinces us that these starlike regions are general, with the degree n case having $2n$ branches going off to infinity, and always containing the real axis. By moving to polar coordinates we can prove that this starlike region always contains the unit circle, as well, because $\Im(z^n) = \Im(\exp(n\ln z)) = r^n\sin(n\theta)$ which will be less than 1 and hence less than π if $r \leq 1$. We have shown

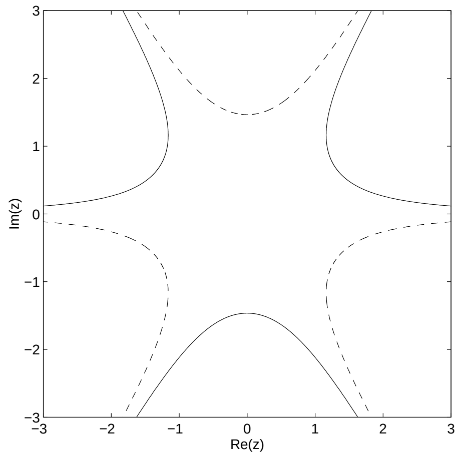

Figure 1: The clearcut region for z^3. Within this star (which contains the unit circle and the rays $\sin 3\theta = 0$) we have $\ln \exp z^3 = z^3$.

that $\mathcal{K}(z^n) = 0$ if $|z| < 1$, for any n, and also that the clearcut region for z^n contains the rays $\theta = k\pi/n$.

6.2 **clearcut**$(\exp z)$

To find the clearcut region for $\exp z$ note that $z = x + iy$ and so $\exp z = \exp(x)\cos x + i\exp(x)\sin(y)$ and so $-\pi < \exp(x)\sin y \le \pi$. Note that if we change y to $y + 2\pi m$ for any integer m this makes no difference; thus the clearcut region will be periodic with period 2π. Hence the boundary curves will be $x = \ln(\pi/\sin y)$ and $x = \ln(-\pi/\sin y)$ and $0 < y < \pi$ works in the first and $-\pi < y < 0$ in the second completes a period. The minimum x-value of these curves occurs at $y = \pm\pi/2$ when we have $x = \ln \pi$. The curves are drawn, with closure information, in Figure 2. The clearcut region of $\exp z$ is the region to the left of the 'fingers'.

6.3 **clearcut**$(\sin z)$

The clearcut region for $\sin z$ is determined by $-\pi < \Im(\sin z) \le \pi$ as before. This gives $-\pi < \cos(x)\sinh(y) \le \pi$, and this region is plotted in Figure 3. We see that if we know that $-\eta < y < \eta$ where $\eta = \sinh^{-1}(\pi) \approx 1.862\ldots$, then $\mathcal{K}(\sin z) = 0$.

6.4 **clearcut**$(\tan z)$

To find the clearcut region for $\tan z$ we proceed as before. The imaginary part of $\tan(x + iy)$ is $\sinh(2y)/(\cos(2x) + \cosh(2y))$, and for a large percentage of x-values this is not larger than 1, in magnitude, and hence $\mathcal{K}(\tan z) = 0$ most of the time. However, for some values of x, for which $\cos 2x$ is negative and nearly -1, then some values of y will give nonzero $\mathcal{K}$.

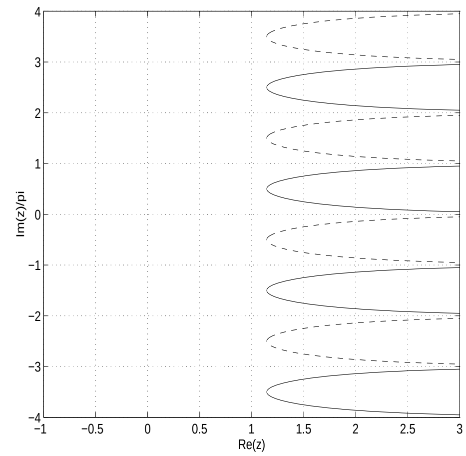

Figure 2: The clearcut region for $\exp z$. Everywhere to the left of the 'fingers' we have $\ln \exp \exp z = \exp z$. The region is closed on the solid lines, open on the dashed lines.

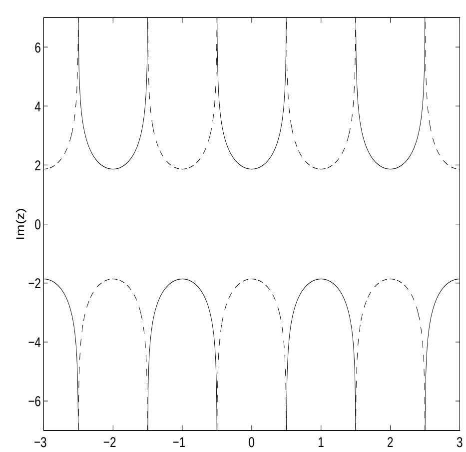

Figure 3: The clearcut region for $\sin z$, with closure as indicated. $\mathcal{K}(\sin z) = 0$ for real z, and also near $\Re(z) = (2m+1)\pi/2$. The clearcut region for $\cos z$ is similar but phase-shifted.

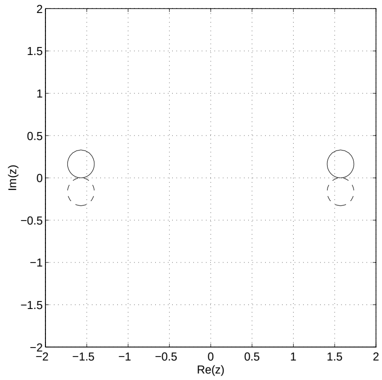

Figure 4: The clearcut region for $\tan z$ is everywhere outside the small isolas which occur near the singularities of $\tan z$. The isolas are periodic in x with period π.

To be precise, if x is between $\pi/2 - (1/2)\sin^{-1}(1/\pi)$ and $\pi/2 + (1/2)\sin^{-1}(1/\pi)$ then there are two values of y, namely $y_\pm$ where $2y_\pm =$

$$\ln\left(\frac{-\pi\cos(2x) \pm \sqrt{\pi^2(\cos(2x))^2 + 1 - \pi^2}}{\pi - 1}\right)$$

where $\Im(\tan(x+iy)) = \pi$. For $y_- \le y \le y_+$ we have $\mathcal{K}(\tan(x+iy)) \ne 0$ (note $y_- > 0$ for $x \ne \pi/2$). This gives a small island or isola near $x = \pi/2$ in which we cannot simplify $\ln\exp\tan z$ to $\tan z$. There is a symmetric isola below the x-axis, open instead of closed because of our choice of closure for $\ln z$, for which the same is true. These isolas are periodic in x with period π. See Figure 4.

6.5 **clearcut**($\sin^{-1} z$)

The clearcut region for $\sin^{-1} z$ is best found by writing down parametric equations for the boundary, namely

$$\sin^{-1} z = t + ip$$

where $p = \pm\pi$. This gives $z = \sin(t+ip) = \sin(t)\cosh(p) + i\sinh(p)\cos(t)$ and so the equations of the boundary region are

$$\begin{aligned} x &= \cosh(\pi)\sin t \\ y &= \pm\sinh(\pi)\cos t \end{aligned}$$

which describes an ellipse in $\mathbb{C}$. Since $\sinh\pi$ and $\cosh\pi$ are both approximately 11.5, this elliptical region looks nearly circular. In the interior of this elliptical region, we have $\mathcal{K}(\sin^{-1} z) = 0$. It is interesting to note that the branch cuts for $\sin^{-1} z$ penetrate the ellipse to $(\pm 1, 0)$, but that along the branch cuts $\mathcal{K}(\sin^{-1} z)$ is still zero iff z is inside the ellipse.

The clearcut region for $\cos^{-1} z$ is bounded by exactly the same ellipse; all that changes is the parameterization of the boundary in the above process. This means that $\ln\exp\sin^{-1} z$ and $\ln\exp\cos^{-1} z$ can both be simplified in the interior of the same z-region. We find, however, that the top of the ellipsoidal disk is closed for $\mathcal{K}(\sin^{-1} z)$ while the bottom is open, and this is reversed for $\mathcal{K}(\cos^{-1} z)$, so care must be taken on the boundary.

6.6 **clearcut**($\tan^{-1} z$)

Both approaches (direct and parametric) work in this case. By Maple, the imaginary part of $\tan^{-1}(x+iy)$ is $\ln((x^2+(y+1)^2)/(x^2+(y-1)^2))/4$, which can be interpreted as half the logarithm of the ratio of the distance from z to the branch point at $-i$ to the distance to the branch point at i. This ratio will be $\exp(2\pi)$ when the imaginary part of $\tan^{-1} z$ is π and $\exp(-2\pi)$ when the imaginary part is $-\pi$. This occurs on circular loci, centred just above i and below $-i$ (in fact at $\pm\coth(2\pi)i \approx \pm 1.000006975i$), and of radius $r = 1/\sinh(2\pi) \approx 0.0037$.

The clearcut region for $\tan^{-1} z$ is everywhere *outside* these extremely tiny circles[2]. This conclusion is so surprising that it is worthwhile drawing the clearcut region another way. We set $\tan^{-1} z = t + ip$ where $-\pi < p \le \pi$. Then $z = \tan(t+ip)$ and we have parametrically

$$\begin{aligned} x &= \frac{\sin 2t}{\cosh 2p + \cos 2t} \\ y &= \frac{\sinh 2p}{\cosh 2p + \cos 2t} \end{aligned}$$

which when plotted with $p = \pi$ gives us the upper circle described above, and with $p = -\pi$ gives us the lower circle.

Thus it is nearly always true that $\ln\exp\tan^{-1} z = \tan^{-1} z$. One would make an error in doing this simplification only if later z was specialized to be within roughly 0.0037 of the branch points at $\pm i$.

7 Concluding Remarks

We have introduced a new mathematical function, the unwinding number, to make computer algebra in $\mathbb{C}$ simpler. This function allows us to give correct rules for manipulating formulas which contain variables taking complex values. Of the many definitions possible, we believe that the definition of $\mathcal{K}(z)$ is the simplest.

The principal benefits of $\mathcal{K}$ include

1. that it allows encapsulation of geometric information for use by 'assume' systems or for human intervention, and

[2]RMC believes he remembers W. Kahan mentioning this fact in conversation some years ago.

2. that it gives a precise definition for 'simplify/symbolic' in Maple or other CAS, that can be used for provisos: one rewrites an expression using $\mathcal{K}$, one uses whatever assumptions one has to evaluate as many instances of $\mathcal{K}$ as possible, and than one sets to zero whatever unwinding numbers are left. The proviso for the result is then just the unwinding numbers that had been set to zero.

The principal *disadvantages* in using $\mathcal{K}$ in a computer algebra system include

1. that rewrite rules using $\mathcal{K}$ essentially double the size of the printed output (though not the DAG), giving answers of the form $Y + 2\pi i\mathcal{K}(Y)$, and
2. that the rules for removal of $\mathcal{K}$ are essentially geometric and need decisions to be taken on the basis of where its arguments are in $\mathbb{C}$.

Automatic geometric reasoning with elementary functions is not well understood yet, and indeed this may prove to be a "grand challenge" to symbolic computation systems, with many other possible applications. Perhaps we may turn this disadvantage of $\mathcal{K}$ into a stimulant for development in this area.

More work needs to be done before this function can be properly implemented. We invite discussion of this function, and in particular we invite discussions containing trial implementations in real computer algebra systems. The primary purpose of this present article is to help to get people *used to* the idea of the unwinding number; of course such a psychological adjustment—to learn to think of $\mathcal{K}$ as an answer, not a question—is a necessary preliminary to its being used in practice. We invite you to check the results in this paper, and to draw some clearcut regions for yourselves (e.g. for $\sqrt{1-z^2}$ or the hyperbolic functions) to help make that adjustment.

Mathematicians make progress by turning analysis into algebra. We hope that $\mathcal{K}(z)$ will help to turn complex analysis into computer algebra.

Acknowledgements

We have discussed this topic with many colleagues over the years, including Helmer Aslaksen, Sam Dooley. Richard Fateman, Michael Monagan, Charlie Patton, and Al Rich. They read an earlier draft of this column and gave helpful feedback, for which we are grateful.

References

[1] T. M. Apostol, Mathematical Analysis, 2nd ed., Addison-Wesley, 1974.

[2] H. Aslaksen, "Multiple-valued complex functions and computer algebra", this BULLETIN, *pp.* 12–20.

[3] Robert M. Corless, Gaston H. Gonnet, D. E. G. Hare, D. J. Jeffrey, & D. E. Knuth, "On the Lambert W Function", Advances in Computational Mathematics, *in press*.

[4] D. J. Jeffrey, D. E. G. Hare, & Robert M. Corless, "Unwinding the branches of the Lambert W function", *Math. Scientist* **21** 1–7 (1996).

[5] Richard J. Fateman, "Why Computer Algebra Systems Sometimes Can't Solve Simple Equations", this BULLETIN, *pp.* 8–11.

[6] 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.

[7] A. D. Rich & D. J. Jeffrey, "Function Values on Branch Cuts" this BULLETIN, *pp.* 25–27.

[8] C. M. Patton, "A Representation of Branch-Cut Information", this BULLETIN, *pp.* 21–24.