"According to Abramowitz and Stegun" or arccoth needn't be uncouth

Robert M. Corless
David J. Jeffrey
Stephen M. Watt
Ontario Research Centre
for Computer Algebra
www.orcca.on.ca

James H. Davenport*
Dept. Mathematical Sciences
University of Bath
Bath BA2 7AY
England
jhd@maths.bath.ac.uk

Abstract

This paper[1] addresses the definitions in OpenMath of the elementary functions. The original OpenMath definitions, like most other sources, simply cite [2] as the definition. We show that this is not adequate, and propose precise definitions, and explore the relationships between these definitions.

In particular, we introduce the concept of a couth pair of definitions, e.g. of arcsin and arcsinh, and show that the pair arccot and *arccoth* can be couth.

1 Introduction

Definitions of the elementary functions are given in many textbooks and mathematical tables, such as [2, 7, 16]. However, it is a sad fact that these definitions often require a great deal of common sense to interpret them, or, to be blunt, are blatantly self-inconsistent: not just at a finite number of special points, but over half the complex plane or half the real line. This state of definition is insufficient for computer programming of any kind, and particularly so for a formal standard, such as OpenMath [1, 5], intended to allow expressions to be conveyed between systems, with all their semantics intact.

The OpenMath definitions of complex elementary functions (contained in the Content Dictionaries[2] **transc1** and **transc2** [11]) generally state that the elementary functions are "as defined in Abramowitz and Stegun" [2], and therefore implicitly single-valued functions. It is the contention of this paper that this is insufficient, and more needs to be done to define the elementary functions precisely, particularly (but not exclusively) on their branch cuts. There are even implications for the definition of functions on **R**, notably arccot. This requirement for unambiguity affects the definitions of the logarithm, non-integer powers such as square root, and the inverse trigonometric and hyperbolic functions, as well as relations between these functions, and relations between them and the forward functions.

We discuss the implications of accurate translation on the design of the phrasebooks [5] translating between OpenMath and actual systems. Note that OpenMath does not say that one definition is 'right' and another 'wrong': it merely provides a *lingua franca* for passing semantically accurate representations between systems. Semantically correct phrasebooks would deduce that

$$\underbrace{\operatorname{arccot}}_{\text{Maple}} z = \frac{\pi}{2} - \overline{\underbrace{\arctan}_{\text{Derive}} \overline{z}}.$$

Notation. Throughout this paper, we use arccot etc to mean the precise function definitions we are using, and variants are indicated by annotations as in the equation above. Throughout, z and its decorations indicate complex variables, x and y real variables. The symbol $\Im$ denotes the imaginary part, and $\Re$ the real part, of a complex number. The precise OpenMath proposals are listed at the end.

2 The base case: Logarithms

Let us begin with the simplest example, and the one in terms of which we will define the other functions.

The definition of $\ln z$ (and hence that of $z^a = \exp a \ln z$ for $a \notin \mathbf{Z}$). [2, p. 67] gives the branch cut $(-\infty, 0]$, and the rule [2, (4.1.2)] that

$$-\pi < \Im \ln z \leq \pi. \quad (1)$$

This then completely specifies the behaviour of ln: on the branch cut it is closed on the positive imaginary side of the cut, i.e. counter-clockwise continuous in the sense of [10].

What are the consequences of this definition[3]? From the existence of branch cuts, we get the problem of a lack of continuity:

$$\lim_{y \to 0^-} \ln(x + iy) \neq \ln x : \quad (2)$$

*This work was performed while the second author held the Ontario Research Chair in Computer Algebra at the University of Western Ontario. All authors are grateful for the comments of Dave Hare at Waterloo Maple Inc. and of many members of the OpenMath Esprit project.

[1] A previous version was a deliverable of the Esprit OpenMath Project, number 24969.

[2] In the process of being reshuffled to ensure alignment with MathML version 2 [15], but this does not affect the thrust of this paper.

[3] Which we do not contest: it seems that few people today would support the rule one of us (JHD) was taught, viz. that $0 \leq \Im \ln z < 2\pi$. The placement of the branch cut is 'merely' a notational convention, but an important one. If we wanted a function that behaves like ln but with this cut, we could consider $\ln(-1) - \ln(-1/z)$ instead. We note that, until 1925, astronomers placed the branch cut between one day and the next at noon [6, vol. 15 p. 417].

for $x < 0$ the limit is $\ln x - 2\pi i$. Related to this is the fact that

$$\ln \overline{z} \neq \overline{\ln z} \tag{3}$$

on the branch cut: instead $\ln \overline{z} = \overline{\ln z} + 2\pi i$ on the cut. Similarly,

$$\ln \frac{1}{z} \neq -\ln z \tag{4}$$

on the branch cut: instead $\ln \frac{1}{z} = -\ln z + 2\pi i$ on the cut.

Two families of solutions have been mooted to these problems.

- [10] points out that the concept of a "signed zero"[4] [9] (for clarity, we write the positive zero as 0^+ and the negative one as 0^-) can be used to solve problems such as the above, if we say that, for $x < 0$, $\ln(x+0^+i) = \ln(x)+\pi i$ whereas $\ln(x+0^-i) = \ln(x)-\pi i$. Equation (2) then becomes an equality for all x, interpreting the x on the right as $x+0^-i$. Similarly, (3) and (4) become equalities throughout. Attractive though this proposal is (and OpenMath should probably not inhibit systems that do have signed zeros), it does not answer the fundamental question: what to do if the user types $\ln(-1)$.

  A similar idea is proposed in [12], who make the functions two-valued on the branch cuts, so that $\ln(-1) = \pm\pi i$. This has the drawback of not fitting readily with numerical evaluation.

  More importantly, neither scheme addresses problems off the branch cuts.

- [4] points out that most 'equalities' do not hold for the complex logarithm, e.g. $\ln(z^2) \neq 2\ln z$ (try $z = -1$), and its generalisation

  $$\ln(z_1 z_2) \neq \ln z_1 + \ln z_2. \tag{5}$$

  The most fundamental of all non-equalities is $z = \ln \exp z$, with an obvious violation at $z = 2\pi i$. They therefore propose to introduce the *unwinding number* $\mathcal{K}$, defined[5] by

  $$\mathcal{K}(z) = \frac{z - \ln \exp z}{2\pi i} = \left\lceil \frac{\Im z - \pi}{2\pi} \right\rceil \in \mathbf{Z} \tag{6}$$

  (note that the apparently equivalent definition $\lfloor \frac{\Im z + \pi}{2\pi} \rfloor$ differs precisely on the branch cut for ln as applied to $\exp z$). (5) can then be rescued as

  $$\ln(z_1 z_2) = \ln z_1 + \ln z_2 - 2\pi i \mathcal{K}(\ln z_1 + \ln z_2). \tag{7}$$

  Similarly (3) can be rescued as

  $$\ln \overline{z} = \overline{\ln z} - 2\pi i \mathcal{K}(\overline{\ln z}). \tag{8}$$

  Note that, as part of the algebra of $\mathcal{K}$, $\mathcal{K}(\overline{\ln z}) = \mathcal{K}(-\ln z) \neq \mathcal{K}(\ln \frac{1}{z})$.

These problems translate into difficulties with the powering operation, even in cases as apparently simple as square roots. In fact,

$$\sqrt{z_1 z_2} = \sqrt{z_1}\sqrt{z_2}\,(-1)^{\mathcal{K}(\ln z_1 + \ln z_2)}. \tag{9}$$

[4]One could ask why zero should be special and have two values (or four in the Cartesian complex plane). The answer is that all the branch cuts we need to consider are on either the real or imaginary axes, so the side to which the branch cut adheres depends on the sign of the imaginary or real part, including the sign of zero. With sufficient care, this technique can be used for other branch cuts.

[5]Note that the sign convention here is the opposite to that of [4], which defined $\mathcal{K}(z)$ as $\lfloor \frac{\pi - \Im z}{2\pi} \rfloor$: the authors of [4] recanted later to keep the number of -1s occurring in formulae to a minimum.

3 General Principles

There is no 'right' answer to the specialisation of the one-many inverse trigonometric and hyperbolic functions. Here we outline various principles that could be used to justify the definitions of the functions (principles 1–2) and the choices of one branch over another (principles 3–5). It should be noted that we cannot apply these principles to each function in isolation, since the choices affect the validity of many identities between functions (principle 6).

1. *For a trigonometric or hyperbolic function f, it should be the case that*

   $$f(f^{-1}(z)) = z, \tag{10}$$

   at least off the branch cuts, and preferably on it. While this might seem obvious, it does need to be checked. Given arcsin, one might construct $\underbrace{\text{arccos}}_{\text{broken}}$ as follows.

   $$\begin{aligned} \sin t &= \sqrt{1 - \cos^2 t}; \\ t &= \arcsin \sqrt{1 - \cos^2 t} \end{aligned}$$

   so, writing $t = \underbrace{\text{arccos}}_{\text{broken}} u$,

   $$\underbrace{\text{arccos}}_{\text{broken}} u = \arcsin \sqrt{1 - u^2}.$$

   However $\cos \arcsin \sqrt{1 - u^2}$ is not the identity function, taking -1 to 1. In general, formal manipulation is not guaranteed to yield an actual inverse.

   Since the functions f are many–one, asking for the converse $f^{-1}(f(z)) = z$ is impossible. However, we could ask that the domain of validity of

   $$f^{-1}(f(z)) = z \tag{11}$$

   be as 'natural' as possible.

2. *All these functions should be mathematically[6] defined in terms of* ln, *thus inheriting their branch cuts from the chosen branch cut for* ln *(equation 1).*

   This does not define the function uniquely: far from that, but it does ensure that the functions can be unambiguously reduced to logarithms, a bonus for simplification programs. This rule is essentially the same as the "Principal Expression" rule of [10]; the difference is that [10] allows $\sqrt{\ }$ as a constructor. In practice we use it as well, but consider it to be defined by $\sqrt{z} = \exp(\frac{1}{2}\ln z)$.

3. *The choice made for the branch cuts over* $\mathbf{C}$ *should be consistent with the 'well-known' behaviour over* $\mathbf{R}$. This generally implies that branch cuts should avoid the real

[6]This does not imply that it is always right to compute them this way. There may be reasons of efficiency, numerical stability or plain economy (it is wasteful to compute a real arcsin in terms of complex logarithms and square roots) why a numerical, or even symbolic, implementation should be different, but the *semantics* should be those of this definition in terms of logarithms, possibly augmented by exceptional values when the logarithm formula is ill-defined.

axis, passing through infinity if necessary: arctan is a good example.

Again, this doesn't disambiguate completely, and may only be a weak guide if the behaviour over **R** is poorly specified.

4. *As many as possible of the 'well-known' identities should hold, at least off the branch cuts.*

   Even more than item (3) above, this is subjective, but still worth adhering to. The decision as to which identity is 'more fundamental' than another often depends on the application being considered, and can therefore seem very controversial.

5. *Occam's razor.*

   The plethora of trigonometric and hyperbolic functions is due partly to the requirements of mathematical tables and their users. For example, cot is frequently tabulated because, near $\pi/2$, it is much easier to use than a table of tan, in which interpolation is difficult, even though it is logically redundant in tables (and calculators) that contain reciprocals and tangents. In general, we should not define functions in OpenMath by formulae which are subtly different from other formulae. The programming language Pascal [8] takes this principle to extremes: the only trigonometric functions defined are sin, cos and arctan. No hyperbolic functions are defined. However, OpenMath is meant to be usable for encoding the full range of mathematical texts, and these do use functions other than sin, cos and arctan. However, the definitions of versine and haversine should probably be relegated to the `historical` CD.

   There is one exception to this that we will see later: we define arccot (and arccoth) *ab initio* rather than in terms of arctan (or arctanh) in order to prevent remarkable behaviour near zero (which would otherwise be inherited from unremarkable behaviour near infinity).

6. *Couthness*

   Since there are well known relations between the hyperbolic and trigonometric functions, there should be similar ones between their inverse functions. If h is any hyperbolic function, and t the corresponding trigonometric function, we have a relation

$$t(z) = ch(iz) \text{ where } c = \begin{cases} 1 & \cos, \sec \\ i & \cot, \operatorname{cosec} \\ -i & \sin, \tan \end{cases}. \qquad (12)$$

From this it follows *formally* that

$$h^{-1}\left(\frac{1}{c}z'\right) = it^{-1}(z'). \qquad (13)$$

Definition 1 *A choice of branch cuts for h^{-1} and t^{-1} is said to be a* couth *pair of choices if equation (13) holds except possibly at finitely many points.*

Of course, there could exist multiple pairs of definitions for h^{-1} and t^{-1}, each pair being couth in itself.

4 Inverse Trigonometric Functions

In this section, we describe the OpenMath definitions for these functions in terms of logarithms (principle 2), and highlight some of the identities and non-identities that result.

4.1 Definition of $\arcsin z$

[2, page 79] gives the branch cuts[7] $(-\infty, -1)$ and $(1, \infty)$, but does not specify the values on the branch cut. Two obvious identities that we want arcsin to satisfy are

$$\arcsin(-z) = -\arcsin(z) \qquad (14)$$

and

$$\overline{\underbrace{\arcsin}_{\text{wishful}}(z)} = \underbrace{\arcsin}_{\text{wishful}}(\overline{z}). \qquad (15)$$

In fact the second cannot be satisfied on these branch cuts: $\arcsin(2) = \frac{\pi}{2} - 1.3169\ldots i$, so $\underbrace{\arcsin}_{\text{wishful}}(\overline{2})$ would have to be $\frac{\pi}{2} + 1.3169\ldots i$, but $2 = \overline{2}$. As in the case of the logarithm, these problems can be resolved by *signed zeros* [10] or by unwinding numbers [4]. For signed zeros, $2 = 2 + 0^+ i$, and $\overline{2 + 0^+ i} = 2 + 0^- i$, and we are at liberty to define $\arcsin(2 + 0^+ i) = \frac{\pi}{2} - 1.3169i$ but $\arcsin(2 + 0^- i) = \frac{\pi}{2} + 1.3169i$. However, as pointed out above, we are now equivocal on the value of arcsin 2.

We require that, as functions $\mathbf{R} \to \mathbf{R}$, on $[-1, 1]$ $\sin(\arcsin(x))$ is the identity, and $\arcsin(\sin(x))$ is the identity on $[-\pi/2, \pi/2]$ (and cannot be on a greater range, since sin ceases to be injective).

We adopt the suggestion in [10], that arcsin be defined by the expression

$$\arcsin z = -i \ln\left(\sqrt{1 - z^2} + iz\right). \qquad (16)$$

This has the correct behaviour on **R**, and satisfies (14), even on the branch cut; but not (15). The unwinding number solution to (15) is given by the following result.

Lemma 1 $\overline{\arcsin z}$ *can be 'simplified':*

$$\begin{aligned} \overline{\arcsin z} \;=\; & (-1)^{\mathcal{K}(-\ln(1-z^2))} \arcsin \overline{z} \\ & + \pi\mathcal{K}(-\ln(1+z)) - \pi\mathcal{K}(-\ln(1-z)) \end{aligned} \qquad (17)$$

The proof is in Appendix A.

4.2 Definition of $\arccos z$

This has the same branch cuts as arcsin in [2], and therefore the same problems. As functions $\mathbf{R} \to \mathbf{R}$, on $[-1, 1]$ $\cos(\arccos(x))$ is the identity, and $\arccos(\cos(x))$ is the identity on $[0, \pi]$ (and cannot be on a greater range, since cos ceases to be injective).

We adopt the suggestion in [10], defining

$$\arccos(z) = \frac{\pi}{2} - \arcsin(z) = \frac{2}{i} \ln\left(\sqrt{\frac{1+z}{2}} + i\sqrt{\frac{1-z}{2}}\right). \qquad (18)$$

This satisfies $\arccos(-z) = \pi - \arccos(z)$ everywhere, as we might expect from **R**.

[7] In fact, [2] does not specify the closure of the branch cuts except in the case of ln: we have completed the definitions.

Lemma 2 *An alternative formulation is*

$$\arccos(z) = -i \ln \left(z + i\sqrt{1-z^2} \right). \tag{19}$$

This is trivially equal to:

$$-i \ln \left(z + i\sqrt{(1-z)(1+z)} \right)$$

which may be numerically more stable near $z = \pm 1$. This issue, and the choice between (18) and (19), which may depend on the availability of hardware square-root capability, are beyond the scope of this paper. The proof is in Appendix B.

4.3 Definition of arctan z

[2] gives the following branch cuts: $(-i\infty, -i]$ and $[i, i\infty)$. [10] suggests, and we will adopt, the following definition:

$$\frac{1}{2i}\left(\ln(1+iz) - \ln(1-iz)\right). \tag{20}$$

This satisfies

$$\arctan(-z) = -\arctan(z), \tag{21}$$

and is now the definition adopted by Common Lisp [13, p. 309], replacing

$$-i \ln \left((1+iz)\sqrt{1/(1+z^2)} \right),$$

which does not satisfy (21).

With this definition, we have the following relationship (proved in Appendix D) between arcsin and arctan:

$$\arcsin z = \arctan \frac{z}{\sqrt{1-z^2}} + \pi\mathcal{K}(-\ln(1+z)) - \pi\mathcal{K}(-\ln(1-z)) \tag{22}$$

4.4 Definition of arccot z

[2] (page 79) gives the branch cut $[-i, i]$ from the ninth printing on, with the definition

$$\operatorname{arccot} z = \arctan(1/z). \tag{23}$$

However, according to [10][8], earlier printings gave the definition

$$\operatorname{arccot} z = \frac{\pi}{2} - \arctan z. \tag{24}$$

It should be noted that definition (23), regarded as the definition of a purely real function, has a range of $(-\pi/2, \pi/2) \setminus \{0\}$, with a singularity at $z = 0$ (probably intended to be given the value $\pi/2$, but this has to be inferred from [2]), whereas (24) has a range of $(0, \pi)$ and no singularity. The branch cuts for (24) are clearly those for arctan.

There is no clear standard for the definition of arccot as a function $\mathbf{R} \to \mathbf{R}$: the standard shibboleth is to ask for the definition of $\operatorname{arccot}(-1)$, though any other negative number would do. This gives the following results.

[2]	1st printing	$3\pi/4$	inconsistent
[2]	9th printing	$-\pi/4$	
[7]	5th edition	?	inconsistent
[16]	30th edition	$3\pi/4$	inconsistent
Maple	V release 5	$3\pi/4$	
Axiom	2.1	$3\pi/4$	
Mathematica	[14]	$-\pi/4$	
Reduce	3.4.1	$-\pi/4$	in floating point
Matlab	5.3.0	$-\pi/4$	in floating point
Matlab	5.3.0	$3\pi/4$	symbolic toolbox

In the absence of any consensus, we appeal to Occam's razor, and, since $\cot(z) = 1/\tan(z)$, we wish to define $\operatorname{arccot}(z)$ by (23). In fact, we need to be slightly more careful, since this does not define $\operatorname{arccot}(0)$. To solve this problem and have $\operatorname{arccot}(0) = \pi/2$ (i.e. interpreting $1/0$ as $+\infty$ rather than $-\infty$), we follow [10] and define

$$\operatorname{arccot} z = \frac{1}{2i} \ln \left(\frac{z+i}{z-i} \right). \tag{25}$$

The apparently similar equation[9]

$$\underbrace{\operatorname{arccot}}_{[3]} z = \frac{i}{2} \ln \left(\frac{z-i}{z+i} \right)$$

in fact defines $-\operatorname{arccot}(-z)$, which takes the value $-\pi/2$ rather than $\pi/2$ at $z = 0$, but is equal everywhere else real. In fact, the two differ only on the branch cuts.

4.5 Definition of arcsec z

[2] gives the branch cut $(-1, 1)$. The parsimony principle mentioned above would define this in terms of arccos, viz.

$$\operatorname{arcsec}(z) = \arccos(1/z) = -i \ln \left(1/z + i\sqrt{1 - 1/z^2} \right). \tag{26}$$

This satisfies the rule that $\operatorname{arcsec}(-z) = \pi - \operatorname{arcsec}(z)$. The singularity at $z = 0$ is genuine.

4.6 Definition of arccsc z

The branch cuts are as for arcsec. Again, we should define this in terms of arcsin, viz.

$$\operatorname{arccsc}(z) = \arcsin(1/z) = -i \ln \left(i/z + \sqrt{1 - 1/z^2} \right). \tag{27}$$

This satisfies the rule that $\operatorname{arccsc}(-z) = -\operatorname{arccsc}(z)$. Again, the singularity at $z = 0$ is genuine.

5 Inverse Hyperbolic Functions

This section lays out our proposals for the definition of the corresponding inverse hyperbolic functions. Questions of couthness, i.e. how this material relates to that of the previous section, are deferred until the next section.

[8] Atypically, Kahan is guilty of over-simplification here. The definition (equation 4.4.8) was always (23), with the branch cuts as listed. However the claimed range was $0 \le \operatorname{arccot} z \le \pi$ (note the $\le$ rather than $<$), with the formula $\operatorname{arccot}(-z) = \pi - \operatorname{arccot}(z)$, which is algebraically inconsistent with (23) and with (21), which was quoted in all printings. This inconsistency has been reproduced in as recent a source as [16] on pages 465–466.

[9] This, in its equivalent form $\frac{-1}{2i} \ln \frac{iz+1}{iz-1}$, is quoted in [3, p. 579] as the definition of arccot. The authors are grateful to Prof. Pohst for this information.

5.1 Definition of arcsinh z

[2] gives the branch cuts: $(-i\infty, -i)$ and $(i, i\infty)$. We follow [10], which gives the principal expression

$$\operatorname{arcsinh}(z) = \ln\left(z + \sqrt{1+z^2}\right), \tag{28}$$

and this satisfies the symmetry rule:

$$\operatorname{arcsinh}(z) = -\operatorname{arcsinh}(-z)$$

5.2 Definition of arccosh z

[2] gives the branch cut $(-\infty, 1)$: in addition there is a branch point at $z = -1$. We follow [10], which gives the primary expression

$$\operatorname{arccosh}(z) = 2\ln\left(\sqrt{\frac{z+1}{2}} + \sqrt{\frac{z-1}{2}}\right). \tag{29}$$

5.3 Definition of arctanh z

[2] gives the branch cuts: $(-\infty, -1]$ and $[1, \infty)$. We follow [10], which gives the principal expression

$$\operatorname{arctanh}(z) = \frac{1}{2}\left(\ln(1+z) - \ln(1-z)\right). \tag{30}$$

This satisfies $\operatorname{arctanh}(z) = -\operatorname{arctanh}(-z)$.

5.4 Definition of arccoth z

[2] gives the branch cut $[-1, 1]$. We define arccoth by

$$\operatorname{arccoth} z = \frac{1}{2}\ln\left(\frac{z+1}{z-1}\right). \tag{31}$$

The expression derived from that of arctanh would be

$$\underbrace{\operatorname{arccoth}}_{\text{tanh-based}}(z) = \frac{1}{2}\left(\ln(1+\frac{1}{z}) - \ln(1-\frac{1}{z})\right). \tag{32}$$

However, this has a jump discontinuity in the imaginary part along the real axis, corresponding to an inconsistency in the imaginary part of arctanh near $\pm\infty$, with $\Im \underbrace{\operatorname{arccoth}}_{\text{tanh-based}} z = \pi/2$ for real $z < 0$, and $-\pi/2$ for $z > 0$. An apparently similar formulation would be

$$\operatorname{arccoth}(z) = \frac{1}{2}\left(\ln(-1-z) - \ln(1-z)\right). \tag{33}$$

Superficially, this has a branch cut on $(-\infty, 1]$, but in fact the branch cuts of the two logarithms cancel on $(-\infty, -1)$, so that the branch cut is in fact $[-1, 1]$. This satisfies $\operatorname{arccoth}(-z) = -\operatorname{arccoth}(z)$ except at $z = 0$.

Lemma 3 *Equations (31) and (33) define the same function.*

The proof is in Appendix C.

5.5 Definition of arcsech z

[2] gives the branch cuts: $(-\infty, 0]$ and $(1, \infty)$. Definition in terms of the formula for arccosh would give

$$\operatorname{arcsech}(z) = 2\ln\left(\sqrt{\frac{z+1}{2z}} + \sqrt{\frac{1-z}{2z}}\right). \tag{34}$$

The singularity at $z = 0$ is genuine.

5.6 Definition of arccsch z

[2] gives the branch cut $(-1, 1)$. Definition in terms of the formula for arcsinh would give

$$\operatorname{arccsch}(z) = \ln\left(\frac{1}{z} + \sqrt{1 + \left(\frac{1}{z}\right)^2}\right), \tag{35}$$

This also has a singularity at $z = 0$.

6 How Couth are We?

In this section, we examine how the various choices we have outlined in the previous two sections fit in with the definition of couthness given above.

6.1 Couthness of arcsin / arcsinh

In this case, equation(13) translates into

$$\operatorname{arcsinh}(iz) \stackrel{?}{=} i\arcsin(z). \tag{36}$$

Substituting in equations (16) and (28) gives

$$\ln\left(iz + \sqrt{1+(iz)^2}\right) \stackrel{?}{=} i\left(-i\ln\left(\sqrt{1-z^2} + iz\right)\right),$$

which is patently true. So these two definitions are couth.

6.2 Couthness of arccos / arccosh

In this case, equation(13) translates into

$$\operatorname{arccosh}(z) \stackrel{?}{=} i\arccos(z). \tag{37}$$

Substituting in equations (18) and (29) gives

$$2\ln\left(\sqrt{\frac{z+1}{2}} + \sqrt{\frac{z-1}{2}}\right) \stackrel{?}{=} i\frac{2}{i}\ln\left(\sqrt{\frac{1+z}{2}} + i\sqrt{\frac{1-z}{2}}\right).$$

The difference between the two is that the left-hand side has $\sqrt{\frac{z-1}{2}}$, whereas the right-hand side has $i\sqrt{\frac{1-z}{2}}$. Now, a consequence of the branch cut (1) for ln is that every square root has its argument in the range $-\frac{\pi}{2} < \arg \le \frac{\pi}{2}$, so these two cannot be equal everywhere. In fact, they are equal when $\arg\left(\sqrt{\frac{1-z}{2}}\right) \le 0$, which simplifies to $\arg(1-z) \le 0$, i.e. when $\Im z > 0$ or $\Im z = 0$ and $\Re(1-z) \ge 0$, i.e. $\Re z \le 1$. This region is the upper half-plane, including the real axis for $\Re z \le 1$.

6.3 Couthness of arctan / arctanh

In this case, equation(13) translates into

$$\operatorname{arctanh}(iz) \stackrel{?}{=} i\arctan(z). \tag{38}$$

Substituting in equations (20) and (30) gives

$$\frac{1}{2}\left(\ln(1+iz) - \ln(1-iz)\right) \stackrel{?}{=} i\left(\frac{1}{2i}\left(\ln(1+iz) - \ln(1-iz)\right)\right)$$

These are clearly equal everywhere.

6.4 Couthness of arccot / arccoth

In this case, equation(13) translates into

$$\operatorname{arccoth}(-iz) \stackrel{?}{=} i \operatorname{arccot}(z). \tag{39}$$

Substituting in equations (25) and (31) gives

$$\frac{1}{2}\ln\left(\frac{-iz+1}{-iz-1}\right) \stackrel{?}{=} i\left(\frac{1}{2i}\ln\left(\frac{z+i}{z-i}\right)\right).$$

Since the inputs[10] to the logarithms are equal, this is trivial.

6.5 Couthness of arcsec / arcsech

In this case, equation(13) translates into

$$\operatorname{arcsech}(z) \stackrel{?}{=} i \operatorname{arcsec}(z). \tag{40}$$

Substituting in equations (26) and (34) gives

$$2\ln\left(\sqrt{\frac{z+1}{2z}}+\sqrt{\frac{1-z}{2z}}\right) \stackrel{?}{=} \ln\left(1/z+i\sqrt{1-1/z^2}\right). \tag{41}$$

At $z=\frac{1}{2}$, this becomes

$$2\ln\left(\sqrt{\frac{3}{2}}+\sqrt{\frac{1}{2}}\right) \stackrel{?}{=} \ln\left(2+i\sqrt{-3}\right):$$

a palpable falsehood. In fact, one side is the negation of the other. Now, lemma 2 shows that

$$\ln\left(z+i\sqrt{1-z^2}\right)=2\ln\left(\sqrt{\frac{z+1}{2}}+i\sqrt{\frac{1-z}{2}}\right).$$

Replacing z by $1/z$ gives

$$\ln\left(\frac{1}{z}+i\sqrt{1-\frac{1}{z^2}}\right)=2\ln\left(\sqrt{\frac{z+1}{2z}}+i\sqrt{\frac{z-1}{2z}}\right).$$

Substituting this into equation (41) gives

$$2\ln\left(\sqrt{\frac{z+1}{2z}}+\sqrt{\frac{1-z}{2z}}\right) \stackrel{?}{=} 2\ln\left(\sqrt{\frac{z+1}{2z}}+i\sqrt{\frac{z-1}{2z}}\right).$$

Since $\ln a = \ln b$ if and only if $a=b$, this reduces to $\sqrt{\frac{1-z}{2z}} \stackrel{?}{=} i\sqrt{\frac{z-1}{2z}}$. These are equal when $\arg\sqrt{\frac{z-1}{2z}} \le 0$, equivalent to $\arg(1-1/z) \le 0$ This is true precisely when $\Im z < 0$ or $\Im z = 0$ and $\Re z \ge 1$. Hence the pair arcsec / arcsech is not couth.

6.6 Couthness of arccsc / arccsch

In this case, equation(13) translates into

$$\operatorname{arccsch}(-iz) \stackrel{?}{=} i \operatorname{arccsc}(z). \tag{42}$$

Substituting in equations (27) and (35) gives

$$\ln\left(\frac{1}{-iz}+\sqrt{1+\left(\frac{1}{-iz}\right)^2}\right) \stackrel{?}{=} \ln\left(\frac{i}{z}+\sqrt{1-\frac{1}{z^2}}\right).$$

These are patently equal.

[10] It is tempting to a computer scientist to write "arguments of", but that way lies linguistic confusion.

7 Implications for Phrase-book Writers

An OpenMath phrase-book is actually a piece of software that translates between the semantics of OpenMath (as defined in the Content Dictionaries) and the semantics of a particular application, as well as simply translating names. Of course, life for the phrase-book writer is simplest if the semantics of the application are the same as those of OpenMath, but this will not always be the case.

A classic example of this would be the translation of arccot between OpenMath, whose semantics we propose to define by (25), and a system, such as Maple or Axiom, where the semantics of arccot are defined by (24).

Application→OpenMath $\quad \operatorname{arccot} z \mapsto \frac{\pi}{2} - \arctan z$

OpenMath→Application $\quad \operatorname{arccot} z \mapsto \arctan(1/z)$

Derive has a different definition of arctan to eliminate the unwinding numbers from (22), so that, for Derive, $\arcsin(z) = \underbrace{\arctan}_{\text{Derive}} \frac{z}{\sqrt{1-z^2}}$. This definition can be related to that of OpenMath either via unwinding numbers or via $\underbrace{\arctan}_{\text{Derive}}(z) = \overline{\arctan \overline{z}}$. It is often possible to deal with such differences on branch cuts by such a 'double conjugate' representation. This representation has the advantage that the (always legal) rule $\overline{\overline{z}} \to z$ means that the transformation is self-inverse.

8 Proposals

This section lists the concrete suggestions that the authors have for OpenMath.

1. OpenMath should *define* a^b (in the case of non-integer b) via
$$a^b = \exp(b \ln a), \tag{43}$$
rather than the current weasel words: "When the second argument is not an integer care should be taken as to the meaning of this function; however it is here to represent general powering" (`arith1.ocd` [11]).

2. OpenMath should base all its single-valued definitions of the naturally multi-valued elementary functions on the logarithm function. This means that it would be reasonable to define all the forward functions in terms of the exponential function.

3. OpenMath should be parsimonious in terms of the number of (subtly) different concepts it introduces: in particular cot, sec and cosec, and their inverses and hyperbolic analogues should be defined[11] in terms of the other functions and reciprocals.

   The actual proposals for the definitions of functions are summarised below.

[11] One could argue that these functions need not be defined at all. Indeed tan could be replaced by sin / cos etc. This would create problems for renderers, but these are not insuperable, and getting good-looking mathematics out of OpenMath already requires a certain amount of intelligence on the part of the renderers. However, not having these functions would, we argue, actually *increase* the risk of the sort of subtle mis-understanding that this paper is meant to point out (e.g. defining arccot in terms of arctan leaves arccot(0) undefined, and the definition is non-trivial: see the discussion after equation (25)), and makes it harder to write numerically-accurate expressions.

arcsin	(16)	arcsinh	(28)	couth
arccos	(18)	arccosh	(29)	uncouth
arctan	(20)	arctanh	(30)	couth
arccot	(25)	arccoth	(31)	couth
arcsec	(26)	arcsech	(34)	uncouth
arccsc	(27)	arccsch	(35)	couth

4. OpenMath should use the 'unwinding number' formalism for stating any mathematical properties of these functions (and therefore an OpenMath CD should define the unwinding number).

5. OpenMath should not forbid the use of signed zeros, and therefore range restrictions such as (1) have to be relaxed to non-strict inequalities in this case.

6. OpenMath should consider whether it wants (using different symbols, e.g. Log) to represent the multi-valued inverse functions as well as the single-valued ones.

References

[1] Abbott,J.A., Díaz,A. & Sutor,R.S, OpenMath: A Protocol for the Exchange of Mathematical Information. SIGSAM Bulletin **30** (1996) 1 pp. 21–24.

[2] Abramowitz,M. & Stegun,I., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. US Government Printing Office, 1964. 10th Printing December 1972. Note that [11] does not define which printing to use.

[3] Bronstein, Semendjajew & Musiol, *Taschenbuch der Mathematik*. Mühlig in Harri Deutsch Verlag, Frankfurt 1995.

[4] Corless,R.M. & Jeffrey,D.J., The Unwinding Number. SIGSAM Bulletin **30** (1996) 2, pp. 28–35.

[5] Dalmas,S., Gaëtano,M. & Watt,S., An OpenMath 1.0 Implementation. Proc. ISSAC '97 (ed. W. Küchlin), ACM New York, 1997, pp. 241–248.

[6] *Encyclopedia Britannica*, 15th. edition. Encyclopedia Britannica Inc., Chicago etc., 15th ed., 1995 printing.

[7] Gradshteyn,I.S. & Ryzhik,I.M. (ed A. Jeffrey), Table of Integrals, Series and Products. 5th ed., Academic Press, 1994.

[8] IEEE Standard Pascal Computer Programming Language. IEEE Inc., 1983.

[9] IEEE Standard 754 for Binary Floating-Point Arithmetic. IEEE Inc., 1985.

[10] Kahan,W., Branch Cuts for Complex Elementary Functions. *The State of Art in Numerical Analysis* (ed. A. Iserles & M.J.D. Powell), Clarendon Press, Oxford, 1987, pp. 165–211.

[11] The OpenMath Consortium. Draft Content Dictionaries. `http://www.nag.co.uk/projects/OpenMath/corecd`

[12] Rich,A.D. and Jeffrey,D.J., Function evaluation on branch cuts. SIGSAM Bulletin 116(1996).

[13] Steele,G.L.,Jr., *Common LISP: The Language*, 2nd. edition. Digital Press, 1990.

[14] Wolfram,S., *The Mathematica Book*. Wolfram Media/C.U.P., 1999.

[15] World-Wide Web Consortium, Mathematical Markup Language (MathML) 2.0 Draft Specification. W3C Draft, revision of 28 March 2000. `http://www.w3.org/TR/2000/WD-MathML2-20000328`.

[16] Zwillinger,D. (ed.), *CRC Standard Mathematical Tables and Formulae*. 30th. ed., CRC Press, Boca Raton, 1996.

A Proof of Lemma 1

The result to be proved is

Lemma 1 $\overline{\arcsin z}$ *can be 'simplified':*

$$\begin{aligned}\overline{\arcsin z} &= (-1)^{\mathcal{K}(-\ln(1-z^2))}\arcsin\overline{z} \\ &\quad + \pi\mathcal{K}(-\ln(1+z)) - \pi\mathcal{K}(-\ln(1-z))\end{aligned} \quad (17)$$

Notice that the expression $\mathcal{K}(-\ln(1-z^2))$ is the most concise characteristic function for the branch cuts of arcsin available in this notation; the other unwinding numbers specify the positive and negative halves of the cut.

Proof: From the definition (16)

$$\overline{\arcsin z} = \overline{i\ \ln(iz+\sqrt{1-z^2})}\ .$$

Since $\overline{\ln z} = \ln\overline{z} + 2\pi i\mathcal{K}(-\ln z)$, this becomes

$$\overline{\arcsin z} = i\ln\left(\overline{iz+\sqrt{1-z^2}}\right) - 2\pi\mathcal{K}\left(-\ln(iz+\sqrt{1-z^2}\ )\right)$$

The unwinding number is zero because it could only be nonzero when the input to log is real and negative, i.e. $iz+\sqrt{1-z^2} = -e^t$ with $t\in\mathbf{R}$. Re-arranging and squaring, $-z^2+2ize^t+e^{2t} = 1-z^2$, so $z = \frac{1-e^{2t}}{2ie^t}$. So z is purely imaginary, but then the log input is always positive. Thus

$$\overline{\arcsin z} = i\ln\left(-i\overline{z}+\overline{\sqrt{1-z^2}}\right)\ .$$

Now $\overline{z^{1/2}} = \overline{z}^{1/2}(-1)^{\mathcal{K}(-\ln z)}$, and therefore

$$\overline{\arcsin z} = i\ln\left(-i\overline{z}+\sqrt{1-\overline{z}^2}\ (-1)^{\mathcal{K}(-\ln(1-\overline{z}^2))}\right)\ . \quad (44)$$

Although (44) is a complete answer to the simplification, it does not explicitly contain arcsin, which is required. Using $\ln z = -\ln z^{-1} - 2\pi i\mathcal{K}(-\ln z)$, we continue.

$$\begin{aligned}\overline{\arcsin z} &= -i\ln\left[\left(-i\overline{z}+\sqrt{1-z^2}(-1)^{\mathcal{K}(-\ln(1-\overline{z}^2))}\right)^{-1}\right] \\ &\quad +2\pi\mathcal{K}\left(-\ln(-i\overline{z}+(-1)^{\mathcal{K}(-\ln(1-\overline{z}^2))}\sqrt{1-z^2})\right)\end{aligned}$$

As above, the last term is zero, and rationalising the first term gives

$$\overline{\arcsin z} = -i\ln\left(i\overline{z}+\sqrt{1-\overline{z}^2}\ (-1)^{\mathcal{K}(-\ln(1-z^2))}\right) \quad (45)$$

If z is not on the branch cut, then $\mathcal{K}$ is zero and we have immediately $\overline{\arcsin z} = \arcsin\overline{z}$. If z is on the branch cut, and again we wish to see arcsin, then returning to (44) gives

$$\overline{\arcsin z} = i\ln\left(-i\overline{z}-\sqrt{1-\overline{z}^2}\right)$$

Invoking $\ln(-z_1) = \ln z_1 + i\pi - 2\pi i\mathcal{K}(i\pi + \ln z_1)$ gives separately

$$\overline{\arcsin z} = \begin{cases} -\arcsin \overline{z} + \pi & z > 1 \\ -\arcsin \overline{z} - \pi & z < -1 \end{cases}$$

Since $\mathcal{K}(-\ln(1-z)) = -1$ for $z > 1$, and similarly for $z < -1$, the lemma follows.

B Proof of Lemma 2

Lemma 2 *An alternative formulation is*

$$\arccos(z) = -i\ln\left(z + i\sqrt{1-z^2}\right). \tag{19}$$

Proof. Now

$$\left(\sqrt{\frac{1+z}{2}} + i\sqrt{\frac{1-z}{2}}\right)^2 = z + i\sqrt{1-z}\sqrt{1+z} = z + i\sqrt{1-z^2}$$

since the imaginary parts of $1-z$ and $1+z$ have opposite signs. Also $2\ln a = \ln(a^2)$ if $\mathcal{K}(2\ln a) = 0$, so we need only show this last stipulation i.e. that

$$-\frac{\pi}{2} < \arg\left(\sqrt{\frac{1+z}{2}} + i\sqrt{\frac{1-z}{2}}\right) \le \frac{\pi}{2}.$$

This is trivially true at $z = 0$. If it is false, then we have to pass through $\left|\arg\left(\sqrt{\frac{1+z}{2}} + i\sqrt{\frac{1-z}{2}}\right)\right| = \frac{\pi}{2}$, i.e. $\sqrt{\frac{1+z}{2}} + i\sqrt{\frac{1-z}{2}} = it$ for $t \in \mathbf{R}$. Squaring both sides, $z + i\sqrt{1-z^2} = -t^2$, i.e. $(z+t^2)^2 = -(1-z^2)$. Hence $2zt^2 + t^4 = -1$, so $z = -(1+t^4)/2t^2 \le -1$, and in particular is real. On this half-line, the argument in question is $+\pi/2$, which is acceptable. Hence the argument never leaves the desired range, and the lemma is proved.

C Proof of Lemma 3

Lemma 3 *Equations (31) and (33) define the same function.*

Proof. This reduces to

$$\ln\left(\frac{z+1}{z-1}\right) = \ln(-1-z) - \ln(1-z).$$

The left-hand side is clearly equal to $\ln\left(\frac{-1-z}{1-z}\right)$, which seems equal to the right-hand side, and in fact differs by $2\pi i\mathcal{K}(\ln(-1-z) - \ln(1-z))$, by the quotient variant of (7). We look at the curves where $\mathcal{K}$ can change value, i.e. the boundaries of the 'clear-cut region' $\mathcal{K} = 0$:

$$\ln(-1-z) - \ln(1-z) = t \pm i\pi \qquad t \in \mathbf{R}$$

and the branch cuts of the individual logarithms, which are $-1-z = -e^t$ and $1-z = -e^t$ (using $-e^t$ for $t \in \mathbf{R}$ as an encoding of the negative real axis). The last two are $z = e^t - 1$ and $z = e^t + 1$, while the first is $\frac{1+z}{1-z} = e^t$, i.e. $z = \frac{e^t-1}{e^t+1}$. Hence all problematic values are for real $z \ge -1$. The complement of this is a connected region, so a trial at, say, $z = -2$ proves that $\mathcal{K} = 0$ everywhere except on this critical line. On the critical line $z \ge -1$, $\arg\ln(-1-z) = \pi$, and we subtract from this the argument of another real logarithm, which is either 0 or π. In either case $\mathcal{K} = 0$.

D arcsin and arctan

The aim of this section is to prove that

$$\arcsin z = \arctan\frac{z}{\sqrt{1-z^2}} + \pi\mathcal{K}(-\ln(1+z)) - \pi\mathcal{K}(-\ln(1-z)) \tag{22}$$

We start from equations (16) and (20). Then

$$\begin{aligned}
& 2i\ \arctan\frac{z}{\sqrt{1-z^2}} \\
= \ & \ln\left(1 + i\frac{z}{\sqrt{1-z^2}}\right) - \ln\left(1 - i\frac{z}{\sqrt{1-z^2}}\right) \\
= \ & \ln\left([1 + i\frac{z}{\sqrt{1-z^2}}]/[1 - i\frac{z}{\sqrt{1-z^2}}]\right) \\
& + 2\pi i\mathcal{K}\left(\ln(1 + i\frac{z}{\sqrt{1-z^2}}) - \ln(1 - i\frac{z}{\sqrt{1-z^2}})\right) \\
= \ & \ln[iz + \sqrt{1-z^2}]^2 \\
& + 2\pi i\mathcal{K}(\ln(1 + i\frac{z}{\sqrt{1-z^2}}) - \ln(1 - i\frac{z}{\sqrt{1-z^2}})) \\
= \ & 2i\arcsin(z) \\
& - 2\pi i\mathcal{K}\left(2\ln(iz + \sqrt{1-z^2})\right) \\
& + 2\pi i\mathcal{K}\left(\ln(1 + i\frac{z}{\sqrt{1-z^2}}) - \ln(1 - i\frac{z}{\sqrt{1-z^2}})\right)
\end{aligned}$$

The tendency for $\mathcal{K}$ factors to proliferate is clear. To simplify we proceed as follows. Consider first the term

$$\mathcal{K}(2\ln(iz + \sqrt{1-z^2}))\ .$$

For $|z| < 1$, the real part of the logarithm's input is positive and hence $\mathcal{K} = 0$. For $|z| > 1$, we solve for the critical case in which the input to $\mathcal{K}$ is $-\pi$ and find only $z = r\exp(i\pi)$, with $r > 1$. Therefore

$$\mathcal{K}(2\ln(iz + \sqrt{1-z^2})) = \mathcal{K}(-\ln(1+z))\ .$$

Repeating the procedure with

$$\mathcal{K}(\ln(1 + iz/\sqrt{1-z^2}) - \ln(1 - iz/\sqrt{1-z^2}))$$

shows that $\mathcal{K} \ne 0$ only for $z > 1$. Therefore

$$\mathcal{K}(\ln(1+iz/\sqrt{1-z^2}) - \ln(1-iz/\sqrt{1-z^2})) = \mathcal{K}(-\ln(1-z))$$

and so finally we get

$$\arctan\frac{z}{\sqrt{1-z^2}} = \arcsin(z) - \pi\mathcal{K}(-\ln(1+z)) + \pi\mathcal{K}(-\ln(1-z)) \tag{46}$$

and this cannot be simplified further.