Branch Cuts and Formal Methods?

James Davenport
masjhd@bath.ac.uk

University of Bath
Thanks to colleagues in Canada and UK;
Partially supported by EPSRC grant EP/J003247/1 and Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2047/1 – 390685813.

6 August 2024
(updated after the talk)

The (Bourbakist) Theory

In principle, (pure) mathematics is clear about "function".

> *On dit qu'un graphe F est un graphe fonctionnel si, pour tout x, il existe au plus un objet correspondant à x par F (I, p. 40). On dit qu'une correspondance $f = (F, A, B)$ est une fonction si son graphe F est un graphe fonctionnel, et si son ensemble de départ A est égal à son ensemble de définition $\mathrm{pr}_1 F$ [pr_1 is "projection on the first component"]. [Bourbaki, Ensembles]*

So for Bourbaki a function includes the definition of the domain and codomain, and is *total* and *single-valued*. We will write $(F, A, B)_{\mathcal{B}}$ for such a function definition.

Notation

$\mathbf{P}(A)$ denotes the power set of the set A.
For a function f, we write $\mathrm{graph}(f)$ for $\{(x, f(x)) : x \in \mathrm{Domain}(f)\}$ and $\mathrm{graph}(f)^T$ for $\{(f(x), x) : x \in \mathrm{Domain}(f)\}$.

Convention (Generally undocumented)

Where an underspecified object, such as $\sqrt{x}$, occurs more than once in a formula, the same *value, or interpretation, is meant at each occurrence.*

For example, $\sqrt{x} \cdot \frac{1}{\sqrt{x}} = 1$ for non-zero x, even though one might think that one root might be positive and the other negative. More seriously, in the formula for the roots of a cubic $x^3 + bx + c$,

$$\frac{1}{6}\sqrt[3]{-108\,c + 12\,\sqrt{12\,b^3 + 81\,c^2}} - \frac{2b}{\sqrt[3]{-108\,c + 12\,\sqrt{12\,b^3 + 81\,c^2}}},$$

the two occurrences of $\sqrt{12\,b^3 + 81\,c^2}$ are meant to have the same value, similarly $\sqrt[3]{-108\,c + 12\,\sqrt{12\,b^3 + 81\,c^2}}$.

Examples of statements [Dav10]

As statements about equality[1] of functions, we consider these:

$$\sqrt{z-1}\sqrt{z+1} \stackrel{?}{=} \sqrt{z^2-1}. \tag{1}$$

$$\sqrt{1-z}\sqrt{1+z} \stackrel{?}{=} \sqrt{1-z^2}. \tag{2}$$

$$\log z_1 + \log z_2 \stackrel{?}{=} \log z_1 z_2. \tag{3}$$

$$\arctan x + \arctan y \stackrel{?}{=} \arctan\left(\frac{x+y}{1-xy}\right). \tag{4}$$

(1) is valid for $\Re(z) > 0$, also for $\Re(z) = 0$, $\Im(z) > 0$.
(2) is valid everywhere, despite the resemblance to (1).
(3) is valid with $-\pi < \arg(z_1) + \arg(z_2) \leq \pi$.
(4) is valid, even for *real* x, y, only when $xy < 1$.

[1] At least at the moment, this is to be considered as extensional, i.e. do the l.h.s. and r.h.s. give the same results for the same inputs?

(4) is curious: arctan is nice

(as a real-valued function, at least).

$$\arctan x + \arctan y \stackrel{?}{=} \arctan\left(\frac{x+y}{1-xy}\right). \qquad (4)$$

On $\mathbf{R}$, $\frac{-\pi}{2} < \arctan < \frac{\pi}{2}$, so the l.h.s. of (4) is in (all of) $(-\pi, \pi)$ whereas the r.h.s. is only in $(\frac{-\pi}{2}, \frac{\pi}{2})$, so (4) can't be an equality.

In fact there is a "branch cut at infinity", since $\lim_{x\to+\infty} \arctan x = \frac{\pi}{2}$, whereas $\lim_{x\to-\infty} \arctan x = -\frac{\pi}{2}$ and $xy = 1$ therefore falls on this cut of the right-hand side of (4).

This is also the branch cut that many symbolic integrators (used to) fall over.

Setting

Various basic facts

- A 1:1 function f has an inverse function f^{-1}

defined on $\mathrm{Codomain}(f) = \mathrm{Domain}(f^{-1})$.

- A 1:1 continuous function f has a continuous inverse function.
- A 1:1 differentiable function f has a differentiable inverse function.

except when $f' = 0$.

- Similarly a 1:1 analytic function f has an analytic inverse function (except when $f' = 0$).

But all this depends on 1:1, and in general the inverse of a continuous etc. function is multivalued.
One way to see lack of 1:1 is via winding numbers.

Multi-valued functions, e.g. [Car58]

Traditionally written with initial capitals.

- $\sin^{-1}(0) = 0$
- $\mathrm{Sin}^{-1}(0) = \{0 + k\pi : k \in \mathbf{Z}\}$
- $\cos^{-1}(1) = 0$
- $\mathrm{Cos}^{-1}(1) = \{0 + 2k\pi : k \in \mathbf{Z}\}$
- $\sin^{-1}(\frac{1}{2}) = \frac{\pi}{6}$
- $\mathrm{Sin}^{-1}(\frac{1}{2}) = \{\frac{\pi}{6} + 2k\pi : k \in \mathbf{Z}\} \cup \{\frac{5\pi}{6} + 2k\pi : k \in \mathbf{Z}\}$
- $2\,\mathrm{Sin}^{-1}(0) = \{0 + 2k\pi : k \in \mathbf{Z}\}$, but
  $\mathrm{Sin}^{-1}(0) + \mathrm{Sin}^{-1}(0) = \{0 + k\pi : k \in \mathbf{Z}\}$

And $\mathrm{Sin}^{-1}(0) - \mathrm{Sin}^{-1}(0) = \{0 + k\pi : k \in \mathbf{Z}\}$

$\mathrm{Sin}^{-1}(\frac{1}{2}) + \mathrm{Sin}^{-1}(\frac{1}{2}) = \{\frac{2\pi}{6}, \frac{6\pi}{6}, \frac{10\pi}{6}\} + \{2k\pi : k \in \mathbf{Z}\}$, so
$\frac{1}{2}\left(\mathrm{Sin}^{-1}(\frac{1}{2}) + \mathrm{Sin}^{-1}(\frac{1}{2})\right) \ni \frac{3\pi}{6}$, whose sin is not $\frac{1}{2}$.

Possible solutions

- Deal in multi-valued functions. This is difficult (as we have seen), but intellectually honest.
- Use the Riemann surface formalism to underpin the multvalued thinking
- Choose a suitable domain on which f is single-valued, so we can talk about f^{-1}
- But this f^{-1}, on this domain, might not be the same as someone else's f^{-1} on their domain, or on the intersection.
  - In particular, not necessarily the same as a software implementation/table.
- Use a standard definition, which defines a principal domain D, and admits that, as z leaves D, then $f^{-1}(f(z))$ will (probably) have a discontinuity, or "branch cut"
- "The nice thing about standards is that there are so many to choose from". Where applicable,we use [AS64, printing ≥ 9], with behaviour on the branch cut defined by [Kah87].

The branch view: [Cartan1973]

p. 32 “The mapping $y \mapsto e^{iy}$ induces an isomorphism ϕ of the quotient group $\mathbf{R}/2\pi\mathbf{Z}$ on the group $\mathbf{U}$. The inverse isomorphism ϕ^{-1} of $\mathbf{U}$ on $\mathbf{R}/\pi\mathbf{Z}$ associates with any complex number u such that $|u| = 1$, a real number which is defined up to the addition of an integral multiple of 2π; this class of numbers is called the argument of u and is denoted by $\arg u$.” In our notation this is $\left(\mathrm{graph}(\phi)^T, U, \mathbf{R}/2\pi\mathbf{Z}\right)_{\mathcal{B}}$.

p. 33 “We define

$$\log t = \log |t| + i \arg t, \qquad (5)$$

which is a complex number defined only up to addition of an integral multiple of $2\pi i$.” In our notation this is $((5), \mathbf{C}, \mathbf{C}/2\pi i\mathbf{Z})_{\mathcal{B}}$.

p. 33 "For any complex numbers t and t' both $\neq 0$ and for any values of $\log t$, $\log t'$ and $\log tt'$, we have

$$\log tt' = \log t + \log t' \pmod{2\pi i}.\text{''} \qquad (6)$$

p. 33 "So far, we have not defined $\log t$ as a *function* in the proper sense of the word".

p. 61 "$\log z$ has a branch in any simply connected open set which does not contain 0."

So any given branch would be $(G, D, I)_{\mathcal{B}}$, where D is a simply connected open set which does not contain 0, G is a graph obtained from one element of the graph (i.e. a pair $(z, \log(z))$ for some $z \in D$) by analytic continuation, and I is the relevant image set.

Branch Cuts of Elementary Functions [Kah87]

exp / ln $\exp(z+2\pi i)=\exp(z)$. These days the principal domain is generally chosen as $\pi < \Im(z) \le \pi$, which translates to a branch cut for ln along the negative real axis, so that $\ln(-1+\epsilon i) \approx i\pi+\epsilon)$, but $\ln(-1-\epsilon i) \approx -i\pi-\epsilon)$.

tan / atan $\tan(z+\pi)=\tan(z)$. Principal domain is $-\frac{\pi}{2} < \Re(z) \le \frac{\pi}{2}$. This translates into a branch cut for atan on $\{0+iy : |y|>1\}$.

cot / acot $\cot(z+\pi)=\cot(z)$. Today the principal domain is $0 \le \Re(z) < \pi$. This translates into a branch cut for acot on $\{0+iy : |y|<1\}$.

cos / acos $\cos(z+\pi)=\cos(z)=-\cos(z)$. The principal domain is $0 \le \Re(z) < \pi$. This translates into a branch cut for acos on $\{x+0i : |x|>1\}$.

Similarly sec etc. and the hyperbolics sinh etc.

False sense of simplicity

Towards an algorithm (I)

$$\sqrt{z-1}\sqrt{z+1} \stackrel{?}{=} \sqrt{z^2-1}. \tag{1}$$

$$\sqrt{1-z}\sqrt{1+z} \stackrel{?}{=} \sqrt{1-z^2}. \tag{2}$$

(2) is correct but (1) is only partially correct. How can we distinguish? The branch cut of $\sqrt{\ }$ is the negative real axis. Regard $\mathbf{C}(z)$ as $\mathbf{R}(x,y)$. Then the branch cuts of (1) are

$\sqrt{z-1}$ $x < 1, y = 0$

$\sqrt{z+1}$ $x < -1, y = 0$

$\sqrt{z^2-1}$ $2xy = 0; x^2 - y^2 - 1 < 0.$

$[\{-1 < x < 1, y = 0\} \cup \{x = 0, y \text{ free}\}]$

These define semi-algebraic (polynomial equations and inequalities) sets in $\mathbf{R}^2$, so partition $\mathbf{R}^2$ into a finite number of cells (found by Cylindrical Algebraic Decomposition), and analyse each cell C_i (which comes with a sample point s_i) separately.

Towards an algorithm (II)

Q1,. . . ,Q4 are the four quadrants of the Argand diagram ($Q1 = \{x \geq 0, y \geq 0\}$ etc.): the branch cut for $\sqrt{}$ means that $\sqrt{Q2} \subset Q1$ and $\sqrt{Q3} \subset Q4$

$x > 0$ (and not $y - 0, x < 1$)Typical point $z = 2$, and (1) becomes $\sqrt{1}\sqrt{3} \stackrel{?}{=} \sqrt{3}$: correct.

$x < 0; y > 0$ Typical point $z = -1 + i$ and (1) becomes
$$\underbrace{\sqrt{\underbrace{-2+i}_{\text{Q2}}}}_{\text{Q1}} \underbrace{\sqrt{\underbrace{i}_{\text{Q2}}}}_{\text{Q1}} \stackrel{?}{=} \underbrace{\sqrt{\underbrace{(-1+i)^2-1}_{\text{Q3}}}}_{\text{Q4}}, \text{ false}$$

$x < 0; y < 0$ Typical point $z = -1 - i$ and (1) becomes $\sqrt{-2-i}\sqrt{-i} \stackrel{?}{=} \sqrt{(-1-i)^2 - 1}$, also false.

Cuts In principle we need to do similar analysis on these.

Towards an algorithm (III)

Not quite so simple: on each cell, the proposed identity is either *everywhere true* ot *generically false*.

Consider multiplying (1) by $z^2 + 2z + 2$, which vanishes at both $z = -1 + i$ and $z = -1 - i$.

Then this is "accidentally" true at the sample points $s_i = -1 \pm i$, even though false elsewhere in their regions. How do we deal with this?

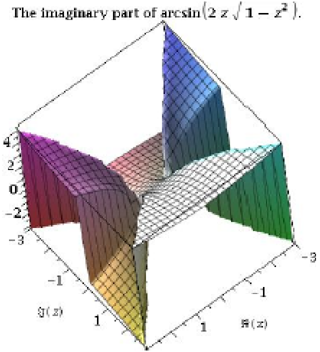
[BBD03] Regard our equation as power series, and use an explicit zero test for these [vdH02].

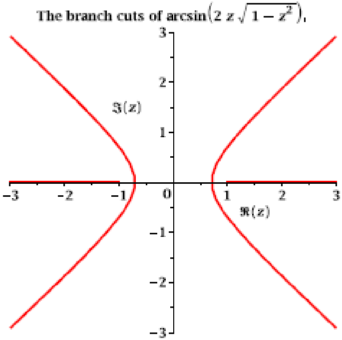
In practice a poly-algorithmic approach is useful [BBDP07], and for branch cuts, we can ask what full-dimensional cell they "adhere" to [BBDP05].

This is implemented in the package `BranchCuts` in Maple: see [EBDW13]. For example given $\operatorname{asin}(2z\sqrt{1-z^2})$, it can produce

$$\{\Im(z)=0, 1<\Re(z)\} \quad \{\Im(z)=\Im(z), \Re(z)=-\tfrac{1}{2}\sqrt{2+4\Im(z)^2}\}$$
$$\{\Im(z)=0, \Re(z)<-1\} \quad \{\Im(z)=\Im(z), \Re(z)=\tfrac{1}{2}\sqrt{2+4\Im(z)^2}\}$$

and the branch cuts on the right (left is a Maple plot).

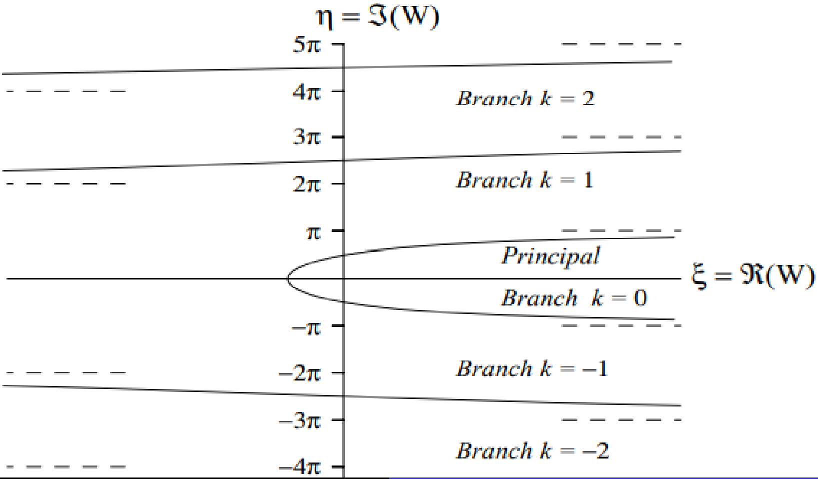




Branch cuts: Lambert W

W is the solution of $W(z)e^{W(z)} = z$, and is not Liouvillian [BCDJ08]. Its branches are more complicated [JHC96].

Figure: Branches of W [JHC96, Figure 2]



Interactions with Proving/Verification

- Can we be more formal than the "proof" I sketched via CAD?
- In an ideal world, that sketch would become a tactic, or possibly a generator of counter-examples.
- And how dependent is this on a "fully verified" CAD?
- Code can be generated from prover output (as in [FM24]), but is that code, with its choice of branch cuts, actually compatible with the prover?
- What if the branch cuts aren't semi-algebraic? As in W.

Interactions with Proving/Verification

- **Lean** I see nothing in [AM24]
- **Isabelle** There's a lot of underpinning stuff around winding numbers in [Gro24], but no branch cuts as such.
- **Rocq** See [Bru11], which treats winding numbers but not branch cuts, and is explicitly "non-constructive".
- **JHD** asked for other input.
- **PVS** NASA have a tool `precision` which, the responder thought, did some of this as well as simple precision checking.

Bibliography I

Jeremy Avigad and Patrick Massot.
Mathematics in Lean Release 0.1.
https://leanprover-community.github.io/mathematics_in_lean/mathematics_in_lean.pdf, 2024.

M. Abramowitz and I. Stegun.
Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing.
US Government Printing Office, 1964.

J.C. Beaumont, R.J. Bradford, and J.H. Davenport.
Better Simplification of Elementary Functions Through Power Series.
In J.R. Sendra, editor, *Proceedings ISSAC 2003*, pages 30–36, 2003.

Bibliography II

J.C. Beaumont, R.J. Bradford, J.H. Davenport, and N. Phisanbut.
Adherence is Better Than Adjacency.
In M. Kauers, editor, *Proceedings ISSAC 2005*, pages 37–44, 2005.

J.C. Beaumont, R.J. Bradford, J.H. Davenport, and N. Phisanbut.
Testing Elementary Function Identities Using CAD.
AAECC, 18:513–543, 2007.

M. Bronstein, R. Corless, J.H. Davenport, and D.J. Jeffrey.
Algebraic properties of the Lambert W function from a result of Rosenstein and Liouville.
J. Integral Transforms and Special Functions, 18:709–712, 2008.

Bibliography III

A. Brunel.
Non-constructive complex analysis in Coq.
https://drops.dagstuhl.de/storage/00lipics/lipics-vol019-types2011/LIPIcs.TYPES.2011.1/LIPIcs.TYPES.2011.1.pdf, 2011.

C. Carathéodory.
Theory of functions of a complex variable.
Chelsea Publ., 1958.

J.H. Davenport.
The Challenges of Multivalued "Functions".
In S. Autexier *et al.*, editor, *Proceedings AISC/Calculemus/MKM 2010*, pages 1–12, 2010.

Bibliography IV

M. England, R. Bradford, J.H. Davenport, and D.J. Wilson.
Understanding Branch Cuts of Expressions.
In J. Carette *et al.*, editor, *Proceedings CICM 2013*, pages 136–151, 2013.

R. Fernández Mir.
Verified Transformations for Convex Programming.
PhD thesis, University of Edinburgh, 2024.

Isabelle Group.
Complex Analysis.
https://isabelle.in.tum.de/library/HOL/HOL-Complex_Analysis/document.pdf, 2024.

D.J. Jeffrey, D.E.G. Hare, and R.M. Corless.
Unwinding the branches of the Lambert W function.
Mathematical Scientist, 21:1–7, 1996.

Bibliography V

W. Kahan.
Branch Cuts for Complex Elementary Functions.
In A. Iserles and M.J.D. Powell, editors, *Proceedings The State of Art in Numerical Analysis*, pages 165–211, 1987.

J. van der Hoeven.
A new Zero-test for Formal Power Series.
In T. Mora, editor, *Proceedings ISSAC 2002*, pages 117–122, 2002.