The Challenges of Multivalued "Functions"

James H. Davenport
J.H.Davenport@bath.ac.uk

Department of Computer Science
University of Bath, Bath BA2 7AY
United Kingdom

Abstract. Although, formally, mathematics is clear that a function is a single-valued object, mathematical practice is looser, particularly with n-th roots and various inverse functions. In this paper, we point out some of the looseness, and ask what the implications are, both for Artificial Intelligence and Symbolic Computation, of these practices. In doing so, we look at the steps necessary to convert existing texts into
(a) rigorous statements
(b) rigorously proved statements.
In particular we ask whether there might be a constant "de Bruijn factor" [18] as we make these texts more formal, and conclude that the answer depends greatly on the interpretation being placed on the symbols.

1 Introduction

The interpretation of "functions" crosses several areas of mathematics, from almost foundational issues to computer science. We endeavour to clarify some of the different uses, starting with the set-theoretic definitions.

Notation 1 $\mathbf{P}(A)$ *denotes the power set of the set A. For a function f, we write* $\mathrm{graph}(f)$ *for* $\{(x, f(x)) : x \in \mathrm{dom}(f)\}$ *and* $\mathrm{graph}(f)^T$ *for* $\{(f(x), x) : x \in \mathrm{dom}(f)\}$.

There is an implicit convention in practically all texts which mention these underspecified objects, which we will not raise further, though there is an exception which is mentioned in section 4.

Convention 1 *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 standard formula for the roots of a cubic $x^3 + bx + c$,

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

the two occurrences of $\sqrt{12\,b^3 + 81\,c^2}$ are meant to have the same value. One could question what is meant by "same formula", and indeed the scope seems in practice to be the entire proof/example/discussion.

Notation 2 *We use the notation* $A \stackrel{?}{=} B$ *to denote what is normally given in the literature as an equality with an* $=$ *sign, but where one of the purposes of this paper is to question the meaning of that, very overloaded [10], symbol.*

We will often need to refer to polar coordinates for the complex plane.

Notation 3 *We write* $\mathbf{C} \equiv X \underset{\text{polar}}{\times} Y$ *for such a representation* $z = re^{i\theta} : r \in X \wedge \theta \in Y$.

As statements about functions, we consider the following exemplars.

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

$$\sqrt{1-z}\sqrt{1+z} \stackrel{?}{=} \sqrt{1-z^2}. \tag{3}$$

$$\log z_1 + \log z_2 \stackrel{?}{=} \log z_1 z_2. \tag{4}$$

$$\arctan x + \arctan y \stackrel{?}{=} \arctan\left(\frac{x+y}{1-xy}\right). \tag{5}$$

For the reader unfamiliar with this topic, we should point out that, with the conventional single-valued interpretations [1], the validity of the formulae is as follows.

(2) is valid for $\Re(z) > 0$, also for $\Re(z) = 0$, $\Im(z) > 0$. It is nevertheless stated as an equation in [11, p. 297] — see section 4.

(3) is valid everywhere, despite the apparent resemblance to (2). One proof is in [6, Lemma 2].

(4) is valid with $-\pi < \arg(z_1) + \arg(z_2) \leq \pi$ — this is [1, 4.1.7].

(5) is valid[1], even for *real* x, y, only when $xy < 1$. This may seem odd, since arctan has no branch cuts over the reals, but 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 (5).

2 The Literature

There is a curious paradox, or at least "abuse of notation" (and terminology) in mathematics to do with the word 'function'.

2.1 The (Bourbakist) Theory

In principle, (pure) mathematics is clear.

> 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"].
>
> [5, p. E.II.13]

[1] Only the multivalued form is given in [1], as 4.4.34.

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. We permit ourselves one abuse of notation, though. The natural domains of definition of analytic functions are simply connected open sets (section 2.3), generally referred to as "$\mathbf{C}^n$ with branch cuts". The table maker, or programmer (section 2.5), abhors "undefined", and extends definitions to the whole of $\mathbf{C}^n$ by making the values on the branch cut 'adhere' [3] to one side or the other, expending a definition from D, a slit version of $\mathbf{C}^n$, to the whole of $\mathbf{C}^n$. Rather than just writing $\mathbf{C}^n$ for the domain, we will explicitly write $\overline{D}$ to indicate that it is an extension of the definition with domain D.

2.2 The multivalued view

Analysts sometimes take a completely multivalued view, as here, discussing our exemplar (4).

> The equation merely states that the sum of one of the (infinitely many) logarithms of z_1 and one of the (infinitely many) logarithms of z_2 can be found among the (infinitely many) logarithms of $z_1 z_2$, and conversely every logarithm of $z_1 z_2$ can be represented as a sum of this kind (with a suitable choice of $\log z_1$ and $\log z_2$).
>
> [7, pp. 259–260] (our notation)

Here we essentially have $\left(\text{graph}(\exp)^T, \mathbf{C}, \mathbf{P}(\mathbf{C})\right)_{\mathcal{B}}$.

Related to this view is the "Riemann surface" view, which can be seen as $\left(\text{graph}(\exp)^T, \mathbf{C}, \mathcal{R}_{\log z}\right)_{\mathcal{B}}$, where $\mathcal{R}_{\log z}$ signifies the Riemann surface corresponding to the function $\log z$. The Riemann surface view is discussed in [6, Section 2.4], which concludes

> Riemann surfaces are a beautiful conceptual scheme, but at the moment they are not computational schemes.

The additional structure imparted by $\mathcal{R}_{\log z}$ (over that of $\mathbf{P}(\mathbf{C})$) is undoubtedly very useful from the theoretical point of view, and provides a global setting for the next, essentially local, view.

2.3 The branch view

Other views may also be found in the analysis literature, for example [8], where one finds the following series of statements.

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 $\left(\text{graph}(\phi)^T, U, \mathbf{R}/2\pi\mathbf{Z}\right)_{\mathcal{B}}$.

p. 33 "We define

$$\log t = \log|t| + i \arg t, \tag{6}$$

which is a complex number defined only up to addition of an integral multiple of $2\pi i$." In our notation this is $((6), \mathbf{C}, \mathbf{C}/2\pi i\mathbf{Z})_{\mathcal{B}}$.

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

$$\log tt' = \log t + \log t' \pmod{2\pi i}.\text{"} \tag{7}$$

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[2] 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.

2.4 An 'Applied' view

Applied mathematics is sometimes less unambiguous.

> ... when we say that $f(x)$ is a function of x in some range of values of x we mean that for every value of x in the range one or more values of $f(x)$ exist. ... It will usually also be required that the function shall be *single-valued*, but not necessarily.
>
> [12, p. 17]

So for these authors, a function might or might not be multivalued.

2.5 The table-maker's point of view

This is essentially also the computer designer's point of view, be it hardware or software. From this point of view, it is necessary to specify how to compute $f(x)$ for any given x, irrespective of any "context", and return a single value, even though, in the text accompanying the tables, we may read "only defined up to multiples of $2\pi i$" or some such.

For the purposes of this discussion, we will use the definitions from [1] (augmented by [9]), but the points apply equally to any other definition of these functions that satisfies the table-maker's criterion of unambiguity.

(2) If we substitute $z = -2$, we obtain $\sqrt{-3}\sqrt{-1} \stackrel{?}{=} \sqrt{3}$, which is false, so the statement is not universally true.

(3) It is impossible to refute this statement.

[2] [8] only defines the concept "branch of log", not a more general definition.

(4) If we take $z_1 = z_2 = -1$, we obtain $\log(-1) + \log(-1) \stackrel{?}{=} \log 1$, i.e. $i\pi + i\pi \stackrel{?}{=} 0$, so the statement is not universally true.

(5) If we take $x = y = \sqrt{3}$, we get $\frac{\pi}{3} + \frac{\pi}{3} \stackrel{?}{=} \frac{-\pi}{3}$, so the statement is not universally true.

2.6 Differential Algebra

A completely different point of view of view is the differential-algebraic one [17]. Here $\sqrt{1-z}$ is an object whose square is $1 - z$, formally definable as w in $\mathbf{C}(z)[w]/(w^2 - (1-z))$. Similarly $\log z$ is a new symbol θ such that $\theta' = 1/z$, and so on for other elementary expressions (we do not say 'functions' here, since they are *not* functions in the Bourbaki sense). From this point of view, our exemplar equations take on a very different allure.

(2) The left-hand side is $vw \in K = \mathbf{C}(z)[v, w]/(v^2 - (z-1), w^2 - (z+1))$, and the right-hand side is $u \in \mathbf{C}(z)[v, w]/(u^2 - (z^2 - 1))$. But to write the equation we have to express $u^2 - (z^2 - 1)$ in K, and it is no longer irreducible, being $(u - vw)(u + vw)$. Depending on which factor we take as the defining polynomial, the equation

$$vw = u \qquad (2')$$

is either true or false (if one were trying to view these as functions, one would say "identically true/false", but that statement has no meaning), and we have to decide which. Once we have decided which, the equation becomes trivially true (or false). The problem is that, with the standard interpretations (which of course takes us *outside* differential algebra), the answer is "it depends on which value of z you have".

(3) The analysis is identical up to the standard interpretations, at which point it transpires that, for the standard interpretations, $vw = u$ is true for all values of z. But, of course, this is what we were trying to prove in the first place.

(4) Here we define θ_1 such that $\frac{\partial \theta_1}{\partial z_1} = \frac{1}{z_1}$ (and $\frac{\partial \theta_1}{\partial z_2} = 0$), θ_2 such that $\frac{\partial \theta_2}{\partial z_2} = \frac{1}{z_2}$ (and $\frac{\partial \theta_2}{\partial z_1} = 0$) and θ_3 such that $\frac{\partial \theta_3}{\partial z_1} = \frac{z_2}{z_1}$ and $\frac{\partial \theta_3}{\partial z_2} = \frac{z_1}{z_2}$. If we then consider $\eta = \theta_1 + \theta_2 - \theta_3$, we see that

$$\frac{\partial \eta}{\partial z_1} = \frac{\partial \eta}{\partial z_2} = 0 \qquad (4'),$$

which implies that η "is a constant".

(5) Again, the difference between the two sides "is a constant".

We have said "is a constant", since the standard definition in differential algebra is that a constant is an object all of whose derivatives are 0. Of course, this is related to the usual definition by the following.

Proposition 1. *A differentiable function $f : \mathbf{C}^n \to \mathbf{C}$, all of whose first derivatives are 0 in a connected open set D, takes a single value throughout D, i.e. is a constant in the usual sense over D.*

The difference between the two can be seen in these "corrected" versions of (4) and (5), where the choice expressions are the "constants".

$$\log z_1 + \log z_2 = \log z_1 z_2 + \begin{cases} 2\pi i & \arg z_1 + \arg z_2 > \pi \\ 0 & -\pi < \arg z_1 + \arg z_2 < \pi \\ -2\pi i & \arg z_1 + \arg z_2 < -\pi \end{cases} \qquad 4''$$

$$\arctan x + \arctan y = \arctan\left(\frac{x+y}{1-xy}\right) + \begin{cases} \pi & xy > 1, x > 0 \\ 0 & xy < 1 \\ -\pi & xy > 1, x < 0 \end{cases} \qquad 5''$$

Equation ($5''$) appears as such, *with* the correction term, as [2, p. 205, ex. 13].

2.7 The pragmatic view

So, which view actually prevails? The answer depends on the context, but it seems to the current author that the view of *most* mathematicians, *most* of the time, is a blend of 2.3 and 2.6. This works because the definitions of differential algebra give rise to power series, and therefore, given "suitable" initial conditions, the expressions of differential algebra can be translated into functions expressed by power series, which "normally" correspond to functions in some open set around those initial conditions.

Whether this is an 'adequate' open set is a more difficult matter — see point 2 in section 5.6.

3 A textbook example

What is one to make of statements such as the following[3]?

$$\int \frac{2^{\sqrt{x}}}{\sqrt{x}} \mathrm{d}x = \frac{2^{1+\sqrt{x}}}{\log 2} + C \qquad (8)$$

We ignore any problems posed by "$\log 2$". The proof given is purely in the setting of section 2.6, despite the fact that the source text is entitled *Calculus*. Translated into that language, we are working in $\mathbf{C}(x, u, \theta)$ where $u^2 = x$ and

$$\theta' = (\theta \log 2)/2u. \qquad (9)$$

(We note that equation (9) implicitly gives effect to Convention 1, in that θ' represents $\left(2^{\sqrt{x}} \log 2\right)/2\sqrt{x}$ where the two occurrences of $\sqrt{x}$ represent the same object.) Similarly the right-hand side is $\frac{2\theta}{\log 2} + C$. Note that, having introduced $2^{\sqrt{x}}$, $2^{1+\sqrt{x}}$ is not legitimate in the language of section 2.6, since the Risch Structure Theorem [16] will tell us that there is a relationship between θ and an η standing for $2^{1+\sqrt{x}}$, viz. that η/θ is constant.

[3] (8) is from [2, p. 189, Example 2].

4 A different point of view

It is possible to take a different approach to these functions, and say, effectively, that "each use of each function symbol means what I mean it to mean at that point". This is completely incompatible with the table-maker's, or the computer's, point of view (section 2.5), but has its adherents, and indeed uses. A classic example of this is given in [11, pp. 294–8], and analysed in [13]. The author considers the Joukowski map $f : z \mapsto \frac{1}{2}\left(z + \frac{1}{z}\right)$ and its inverse $f^{-1} : w \mapsto w + \sqrt{w^2 - 1}$, in two different cases. If we regard these functions as $(f, D, D')_{\mathcal{B}}$ and $\left(f^{-1}, D', D\right)_{\mathcal{B}}$, the cases are as follows.

(i): $D = \{z : |z| > 1\}$. Here $D' = \mathbf{C} \setminus [-1, 1]$. The problem with f^{-1} is interpreting $\sqrt{w^2 - 1}$ so that $|w + \sqrt{w^2 - 1}| > 1$.

(ii): $D = \{z : \Im(Z) > 0\}$. Here $D' = \mathbf{C} \setminus ((-\infty, -1] \cup [1, \infty))$, and the problem with f^{-1} is interpreting $\sqrt{w^2 - 1}$ so that $\Im(w + \sqrt{w^2 - 1}) > 0$.

We require f^{-1} to be injective, which is a problem, since in both cases $w \mapsto w^2$ is not. Hence the author applies (2) formally (though he does not say so explicitly), and writes

$$f^{-1}(w) = w + \sqrt{w + 1}\sqrt{w - 1}. \tag{10}$$

(i) Here he takes both $\sqrt{w + 1}$ and $\sqrt{w - 1}$ to be uses of the square-root function from [1], viz. $\left(\sqrt{\ }, \mathbf{C}, \mathbf{C} \equiv \mathbf{R}^+ \underset{\text{polar}}{\times} (-\frac{\pi}{2}, \frac{\pi}{2}] \cup \{0\}\right)_{\mathcal{B}}$. We should note that this means that $\sqrt{w + 1}\sqrt{w - 1}$ has, at least potentially[4], an argument range of $(-\pi, \pi]$, which is impossible for any single-valued (as in section 2.5) interpretation of $\sqrt{w^2 - 1}$.

(ii) Here he takes $\sqrt{w + 1}$ as before, but $\sqrt{w - 1}$ to be an alternative interpretation: $\left(\sqrt{\ }, \mathbf{C}, \mathbf{C} \equiv \mathbf{R}^+ \underset{\text{polar}}{\times} [0, \pi) \cup \{0\}\right)_{\mathcal{B}}$.

In terms of section 2.1, of course, $(f, D, D')_{\mathcal{B}}$ is a bijection (in either case), so $\left(f^{-1}, D', D\right)_{\mathcal{B}}$ exists, and the question of whether there is a "formula" for it is not in the language.

5 Formalisations of these statements?

Of course, the first question is "which kind of statement are we trying to formalise". This matters in two sense — which of the views in section 2 are we trying to formalise, and are we trying to formalise just the statement, or the statement *and* its proof. The question "which view" seems to be a hard one — when reading a text one often has few clues as to the author's intentions in this area. Nevertheless, let us suppose that the view is given.

[4] This is attained: $w = -2$ gives $-\sqrt{3}$, with argument π, whereas $w = -2 - \epsilon i$ gives a result with argument $-\pi + \epsilon'$.

5.1 The (Bourbakist) Theory

In this view a function is defined by its graph, there is no language of formulae, and the graph of the inverse of a bijective function is the transpose of the graph of the original. Therefore the task of formalising any such statements is the general one of formalising (set-theoretic) mathematical texts.

5.2 The multivalued view

Convention 2 *We use*[5] *capital initial letters to denote the multivalued equivalents of the usual functions, so* $\mathrm{Log}(z) = \{w : \exp(w) = z\}$. *By analogy, we write* $\mathrm{Sqrt}(z) = \{w : w^2 = z\}$.

Here, an expression from the usual mathematical notation, such as (4), becomes, as stated in section 2.2,

$$\begin{aligned} \forall w_3 \in \mathrm{Log}(z_1 z_2) \exists w_1 \in \mathrm{Log}(z_1), w_2 \in \mathrm{Log}(z_2) : w_3 = w_1 w_2 \wedge \\ \forall w_1 \in \mathrm{Log}(z_1), w_2 \in \mathrm{Log}(z_2) \exists w_3 \in \mathrm{Log}(z_1 z_2) : w_3 = w_1 w_2, \end{aligned} \tag{11}$$

There is significant expansion here, and one might be tempted to write

$$\mathrm{Log}(z_1 z_2) = \mathrm{Log}(z_1) + \mathrm{Log}(z_2) \tag{12}$$

using set-theoretic addition and equality of sets, which looks reassuringly like a multivalued version of (4).

However, there are several caveats. One is that, as explained in [6, section 2.3], the correct generalisation of $\log(z^2) = 2\log(z)$ is

$$\mathrm{Log}(z^2) = \mathrm{Log}(z) + \mathrm{Log}(z) \tag{13}$$

(note that Convention 1 does *not* apply here, since we are considering sets of values, rather than merely underspecified values) and *not* $\mathrm{Log}(z^2) = 2\,\mathrm{Log}(z)$ (which, effectively, *would* apply the convention). The second is that not all such equations translate as easily: the multi-valued equivalent of

$$\arcsin(z) \stackrel{?}{=} \arctan\left(\frac{z}{\sqrt{1-z^2}}\right) \tag{14}$$

is in fact

$$\mathrm{Arcsin}(z) \cup \mathrm{Arcsin}(-z) = \mathrm{Arctan}\left(\frac{z}{\mathrm{Sqrt}(1-z^2)}\right). \tag{15}$$

Conclusion 1 *Translating statements about these functions to the multivalued view is not as simple as it seems, and producing* correct *translations can be difficult.*

[5] Note that this convention is often reversed in Francophone countries.

It *might* be possible to define a rewrite with constant expansion (a "de Bruijn factor" in the sense of [18]), e.g. by defining new functions such as $\mathcal{AS}(z) = \text{Arcsin}(z) \cup \text{Arcsin}(-z)$, but to the author's knowledge this has not been done, and would probably be a substantial research project.

It would be tempting to wonder about the difficulties of translating proofs, but, other than his and his colleagues', the author has only seen proofs which work by reduction modulo $2\pi i$, and therefore do not generalise to equations like (15), for which the author only knows his own (unpublished) proof.

As has been said, we see little hope for formalising the more general 'Riemann surfaces' version of this view.

5.3 The branch view

In this view, a function is defined locally, in a simply-connected open set, and the statements made informally in mathematics are true in this interpretation if they are true in the informal sense. The first real problem comes in determining the side conditions, such as "not containing 0". For a *fixed* vocabulary of functions, such as the elementary functions (which can all be derived from exp and log) this can probably be achieved, but when new functions can be introduced, it becomes much harder.

The second problem is to determine what such a suitable open set is, and whether one can be found which is large enough for the mathematical setting envisaged. This is often equivalent to the problem of finding a suitable path, and the challenges are really those of formalising traditional analysis.

5.4 An 'Applied' view

This can hardly be said to be formal, and could therefore be ignored. However, in practice a lot of texts are written this way, and it would seem an interesting challenge to see which statements *can* be formalised. But this is clearly harder than formalising a statement once one knows what the context is.

5.5 The table-maker's point of view

For the elementary functions, the table-maker now has an effective methodology, implicit in [1] and made explicit in [9].

1. Choose a branch cut $\mathcal{X}$ for log, and this defines the value of $\log(z)$ for $z \in \mathbf{C} \setminus \mathcal{X}$ by integration from $\log 1 = 0$.
2. Choose an *adherence rule* to define the value of log on $\mathcal{X}$.
3. For each other function f, choose an expression for $f(x)$ in terms of log. There may be several such choices, and none are perfect.
4. As a consequence of step 3, write down the various simplification rules that f (and other functions) must satisfy.
5. If one is unhappy with these results, return to step 3. **Do not** attempt to rewrite the rules — this leads to inconsistencies, with which tables have been (and alas continue to be) bothered over the years.

Conclusion 2 *In the table-maker's view,* statements *about multi-valued functions, if correct, are the same as usually stated. However, they may require amplification, as in (5") versus (5). At least naïvely, such expansion may be unbounded.*

Proofs are a trickier matter. As far as the author knows, such proofs were generally not published before the days of computer algebra, though the table-makers certainly had intuitive understandings of them, at least as regards real variables. Many such proofs are *ad hoc*, but the dangers in an intuitive approach can be seen in (2) versus (3), where apparently similar equations have very different regions of validity. [4] presents a methodology which, for most[6] such equations is guaranteed to either prove or refute them. However, these methods are expensive, and, as they depend on cylindrical algebraic decomposition, the complexity grows doubly-exponentially with the number of variables. Fortunately, there are very few such identities in practice which require more that two complex variables, but even then the methodology has to treat then as four real variables.

Conclusion 3 *Producing (formal) proofs of such statements is a developing subject, even in the context of a computer algebra system. Converting them into an actual theorem prover is a major challenge. Unfortunately, cylindrical algebraic decomposition, as used here, does not seem to lend itself to being used as an 'oracle' by theorem provers.*

5.6 Differential Algebra

The translation from *well-posed* statements of analysis into this viewpoint is comparatively easy, as seen in section 2.6. There are, however, two significant problems.

1. "Well-posed", in the context of differential algebra, means that every extension that purports to be transcendental really is, *and* introduces no new constants. Hence every simplification rule essentially reduces to "correct up to a constant", and beyond here differential algebra does not help us, as seen with $(5'')$.
2. There is no guarantee that the expressions produced by differential algebra, when interpreted as functions, will be well-behaved. This point has been well-explored elsewhere [14, 15] so we will do no more than draw attention to it here.

5.7 The pragmatic view

The fundamental problem with the pragmatic view is that it is a hybrid, and the formalisms in sections 5.3 and 5.6 are different. Indeed, it is precisely this difference that causes most of the problems in practice. The pragmatist takes

[6] There are some technical limitations on the nesting allowed, but these seem not to occur in practice.

formulae produced in the differential-algebraic viewpoint, and interprets them in the branch viewpoint. In the branch viewpoint, every integral is continuous, but, as in point 2 above, there is no guarantee of this remaining true in the hybrid view unless special care is taken.

Conclusion 4 *The pragmatist's view, while useful, is indeed a hybrid, and great care must be taken when translated from one viewpoint to the other.*

6 Conclusion

The handling of (potentially) multi-valued functions in mathematics is bedevilled by the variety of viewpoints that can be adopted. Texts are not good at stating which viewpoint is being adopted, and trying to deduce which (if any) is being adopted is an untackled AI problem.

The most useful viewpoint, it seems, is based on a dual view (section 2.7) of power-series analysis (section 2.3) and differential algebra (section 2.6): the second being quintessentially symbolic computation. These are ultimately based on different formalisms, and there is no guarantee of a smooth translation of the statements, and there may be a requirement for additional proofs that the translations are indeed valid.

Acknowledgements

The author is grateful to the organisers of AISC for inviting this talk, to Patrizia Gianni and Barry Trager for their hospitality while the talk was conceived, and Russell Bradford, Nalina Phisanbut and Stephen Watt for useful comments.

References

1. M. Abramowitz and I. Stegun. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 9th printing. *US Government Printing Office*, 1964.
2. T.M. Apostol. Calculus, Volume I, 1st edition. *Blaisdell*, 1961.
3. 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.
4. J.C. Beaumont, R.J. Bradford, J.H. Davenport, and N. Phisanbut. Testing Elementary Function Identities Using CAD. *AAECC*, 18:513–543, 2007.
5. N. Bourbaki. Théorie des Ensembles. *Diffusion C.C.L.S.*, 1970.
6. R.J. Bradford, R.M. Corless, J.H. Davenport, D.J. Jeffrey, and S.M. Watt. Reasoning about the Elementary Functions of Complex Analysis. *Annals of Mathematics and Artificial Intelligence*, 36:303–318, 2002.
7. C. Carathéodory. Theory of functions of a complex variable. *Chelsea Publ.*, 1958.
8. H. Cartan. Elementary Theory of Analytic Functions of One or Several Complex Variables. *Addison-Wesley*, 1973.
9. R.M. Corless, J.H. Davenport, D.J. Jeffrey, and S.M. Watt. According to Abramowitz and Stegun. *SIGSAM Bulletin*, 34(2):58–65, 2000.

10. J.H. Davenport. Equality in computer algebra and beyond. *J. Symbolic Comp.*, 34:259–270, 2002.
11. P. Henrici. Applied and Computational Complex Analysis I. *Wiley*, 1974.
12. H. Jeffreys and B. Jeffreys. Methods of Mathematical Physics (3rd edition). *Cambridge University Press*, 1956.
13. W. Kahan. Branch Cuts for Complex Elementary Functions. In A. Iserles and M.J.D. Powell, editors, *The State of Art in Numerical Analysis*, pages 165–211, 1987.
14. D. Lazard and R. Rioboo. Integration of Rational Functions - Rational Computation of the Logarithmic Part. *J. Symbolic Comp.*, 9:113–115, 1990.
15. T. Mulders. A note on subresultants and the Lazard/Rioboo/Trager formula in rational function integration. *J. Symbolic Comp.*, 24:45–50, 1997.
16. R.H. Risch. Algebraic Properties of the Elementary Functions of Analysis. *Amer. J. Math.*, 101:743–759, 1979.
17. J.F. Ritt. Differential Algebra. *American Mathematical Society Colloquium Proceedings vol. XXXIII*, 1950.
18. F. Wiedijk. The De Bruijn Factor. `http://www.cs.kun.nl/~freek/notes/factor.pdf`, 2000.