Prospective, Retrospective, and Formal Division: a contribution to philosophical arithmetic

Jan A. Bergstra
j.a.bergstra@uva.nl, janaldertb@gmail.com
Informatics Institute, University of Amsterdam
Science Park 904, 1098 XH, Amsterdam, The Netherlands

Submitted: 10 May 2022
Revised: 13 October 2022

Abstract
Division by zero is a controversial theme. Why is division by zero a relevant issue and how can this issue be addressed from different perspectives? We use “fracterm” as an abbreviation for “fractional expression.” Three types of occurrence of the division symbol in a fracterm are distinguished: prospective occurrence, retrospective occurrence, and formal occurrence. Mathematics mostly features retrospective occurrences of division, computer programming gives rise to prospective occurrences, and so does automated proof checking. The use of division in an axiom system may indicate the presence of formal occurrences of division.

1 Introduction
The mainstream view on division by zero, say dividing 1 by 0, with one and zero understood as rational or real numbers, amounts to the idea that division by zero is not possible because for no known number $x$ is it the case that $0 \cdot x = 1$. To see the latter one notices that clearly $0 \cdot x = 0$ and that by assumption $0 \neq 1$ so that $0 \cdot x = 1$ cannot hold.

Because this impossibility is apparently self-evident, the nonexistence of $1/0$ is understood to be much more far reaching than the mere observation that a plausible value for $1/0$ cannot be found among the rationals and in reals. Given the supposedly “absolute” irrelevance of $1/0$, the use of the latter notation is often perceived as a sign of a lack of understanding of how professional arithmetic ought to be written.

In most mathematical texts precautions are taken so that for any expression with division as the leading function symbol, it is known to the reader (who is supposedly reading the text in the order of presentation) that the corresponding denominator is non-zero. So the text:
Let \( q \neq 2 \) and consider the expression \( e \) given by \( e \equiv (p + 3)/(q - 2) \), and let ...

is acceptable, at least up to the dots, while the text

Choose a number \( q \), consider the expression \( e \equiv (p + 3)/(q - 2) \),
then we know that \( q \neq 2 \) and moreover let ....

is problematic because at the position in the text where \( e \) is introduced it is not known that \( q - 2 \neq 0 \).

Beyond these explanations the mainstream position seems to imply that even pondering what the value of \( 1/0 \) might be shows a lack of understanding of the topic. A student is soon supposed to grasp the enormity of even contemplating such options.

I have taken the phrase “philosophical arithmetic” from [1] in order to propose a larger theme in which this paper may fit. Although [1] is a practical exposition which nowadays would rather fit under the heading of applied arithmetic, I believe that philosophical arithmetic provides a suitable classification of the work in this paper, as well as for various existing papers on division by zero, including work on number systems with unconventional peripheral numbers.

2 Story lines about Division by Zero

Let us assume that a person \( A \) asks mathematician \( B \) the following question: What is \( 1/0 \)?

What replies by \( B \) can be imagined, which replies are plausible, which replies are relevant and meaningful? I will first sketch four different lines along which \( B \) may respond to this question, depending on \( B \)’s views on the matter. These are four out of a plurality of story lines on division by zero. I will discuss three patterns of response that \( B \) may engage in, in response to the stated question. The order of presentation follows (my view on) the order of relevance of the different views.

2.1 No-nonsense approach: \( 1/0 \) is a fracterm without a value

1. “What is \( 1/0 \)?” is an important question which has bothered mathematicians and philosophers alike for centuries.

2. To begin with \( 1/0 \) is a fracterm. A fracterm is an expression with division as the leading function symbol. A fracterm consists of a numerator, in this case \( 1 \) and a denominator, in this case \( 0 \), and a function symbol, in this case \( / \) for division.

3. A fracterm may or may not have a value, that is, it may or may not denote a number. For the fracterm \( 1/0 \) initially we do not know whether or not it has a value.

4. The idea is that \( a/b \) denotes a number \( c \) with the property that \( a = b \cdot c \). In school we will not come across any number \( x \) with the property that \( 1 = 0 \cdot x \). In fact we will have \( 0 \cdot x = 0 \) for all numbers
that we meet and, we have \( 0 \neq 1 \), so that for all \( x \) we get to know in school it will be the case that \( 0 \cdot x \neq 1 \).

5. In school mathematics, and in fact in most of applied and pure mathematics, we will work as if \( 1/0 \) has no meaning (is non-denoting). Moreover it is assumed in general that \( 0 \cdot x = 0 \) is a fundamental axiom.

6. The argument in item 4 is inconclusive in the case of the fracterm \( 0/0 \) which might be given value 0. Now, however, we find that any \( c \) which satisfies \( 0 \cdot c = 0 \) (that is any number \( c \) which will be discussed in school) can be taken for the meaning of \( 0/0 \). Apparently a form of non-determinism arises in the case of \( 0/0 \) which is unfortunate. In view of this non-determinism we will work as if \( 0/0 \) is non-denoting as well.

7. Nevertheless some complications arise, for instance one may contemplate the equation \( 1/0 = 1/0 \). Faced with this “equation” one may contemplate various options:
   (i) \( 1/0 = 1/0 \) is not a legal equation because it contains one or more expressions (in fact two) that have no value, whether or not it is true is not an issue therefore,
   (ii) \( 1/0 = 1/0 \) can be accepted as an equation, that is as a legal text, but it is assigned no truth value.
   (iii) \( 1/0 = 1/0 \) is accepted as an equation, and in addition it is considered to be wrong, assuming that about non-existing entities no true assertions can be made.

8. Making up one’s mind for choosing one of the three options mentioned above is not a matter of arithmetic, it is a matter of theory of language, of philosophy, or of logic. Not engaging in certain judgements, rather than experimenting with debatable semantic options by itself counts as a credible expression of knowledge of arithmetic.

The ambivalence of these options indicates a difficulty in providing a convincing account of the validity and meaning of say \( 1/0 = 1/0 \). But for the non-nonsense approach to division by zero that difficulty is irrelevant just as much as the existence of a plurality of accounts of the liar paradox is irrelevant:

**Claim 2.1** Just as a user of natural language need not be aware of a meaningful account of the logic of the liar paradox, a user of the language of school arithmetic need not be aware of an account of the logic of various extreme (in the sense of very uncommon) forms of expression in said language.

Although the no-nonsense approach seems to be self-explanatory, there is an intrinsic difficulty to it. The basic idea is that an author can use a fracterm \( p/q \) in a text (or any other form of discourse) whenever it is known that \( q \) is non-zero. That raises the questions: (i) on the basis of which data is this fact known (in particular which data from what context are used for obtaining that knowledge), (ii) what is the argument that allows us to infer \( q \neq 0 \) from said data, and (iii) to whom are these data and the corresponding implication of \( q \neq 0 \) known?
Claim 2.2 When adopting the no-nonsense approach syntax and semantics become essentially intertwined: an occurrence of the fracterm \( \frac{p}{q} \) in a context \( C[-] \) is acceptable if and only if it is known (by reading \( C[p/q] \) from left to right) when reading \( p/q \) that \( q \neq 0 \). Thus: whether or not a text, say \( C[p/q] \) is unproblematic (qua text) depends on the value of \( q \).

Below I will discuss what I call Dedekind’s workshop. That is the set theoretic toolkit which mathematicians use, or claim to use, when making mental constructions of mathematical realities. Such realities are understood as being “perfect” by construction, while the language(s) and notations we use to speak and argue about such realities may be defective. Mathematicians prefer not to contemplate phenomena which they perceive as coming about from the choice of language and notation, while they prefer to focus on the underlying reality (of mental constructions) which is supposed to be language independent to a very high degree. The no-nonsense view regards all issues about division by zero as related to language and notation (that is “nonsense”) and as devoid of content in relation to the realities (that is mathematics proper) that they manufactured, during the last say 150 years, in Dedekind’s workshop.

Claim 2.3 Implicit in the no-nonsense approach is an underlying notion of legality for texts which governs which texts can be written and read and which texts are better dismissed. For instance the question “let \( x = 0 \), is \( \frac{2}{x} > 3 \)” is better dismissed, and as a text it is for that reason illegal. Formalising the notion of legality is against the spirit of the no-nonsense approach: the illegal utterances being “nonsense” these are not worth systematic scrutiny. It is the rule rather than the exception with any (self-portrayed) no-nonsense approach (to whatever theme) that the question “what is nonsense” is not paid much attention.

2.2 Transrationals and rejection of the law \( 0 \cdot x = 0 \)

A second story line adopts the idea of transrationals and transreals. In these versions of arithmetic the presence of a positive infinite value \( \infty \) and its opposite, a negative infinite value \( -\infty \) is adopted. These are so-called peripheral numbers. Besides both infinite peripherals a third one is needed, named \( \Phi \) in the original work on transrationals [2], and the identity \( \infty \cdot 0 = (-\infty) \cdot 0 = \Phi \neq 0 \) is adopted.

1. Concerning multiplication with 0 transrationals adopt \( 0 \cdot \Phi = \Phi \), so that each of the three peripherals defeat \( 0 \cdot x = 0 \).
2. In transrationals and transreals one finds: \( 1/0 = \infty \), as well as \( 1/\infty = 0 \).
3. Although working with \( \infty, -\infty \) and \( \Phi \) may be unfamiliar to most readers, the idea is quite plausible. Transrationals constitute an abstract model similar to the IEEE 754 floating-point arithmetic standard.
4. Almost all of computing is nowadays carried out with an arithmetic where \( 1/0 = +\text{Inf} \) (the symbol \( \infty \) is uncommon for computer output) and \( (-1)/0 = -\text{Inf} \). The suggestion to have calculation or-
organised in that manner was originally made by computer pioneer Konrad Zuse around 1939 (see [4]).

Given the ubiquitous practice with the use of the peripheral number $+\text{Inf}$ in computing there has been remarkably little theoretical work on that matter. The main approach in this direction, named transreals, has been pioneered by James Anderson and his colleagues in a sequence of papers of which we mention [2, 18] and [19]. For further work in this direction see [14].

In fact one may claim that:

**Claim 2.4** \( 1/0 = \infty \) is the most wide-spread “answer” on the question “what is \(1/0)\).

Using \(1/0 = \infty\) the world of fractional expressions becomes quite complicated and no useful standard forms of such expressions have yet been found [8].

Program notations always come with a syntax and it is usually taken for granted that syntactic correctness matters. In the context of computer programming it is highly impractical, not to say computationally unfeasible, to make correct syntax depend on semantics.

For instance mixing logical connectives and arithmetical functions may be disallowed so that “expressions” \((1+2) \land 3, \frac{1+2}{3+4}, 0 \lor 5\) are each rejected. In the no-nonsense approach it is taken for granted that such “expressions” are not taken into consideration, while in the formalised world of a program notation or a notation for mathematical proofs (that, because of size and complexity, will be checked automatically rather than by a human reader), the fact that such expressions are illegal (that is problematic from the perspective of syntax) is made explicit.

**Claim 2.5** If a program notation allows occurrences of division it is not to be expected that precisely those occurrences of division which would be considered acceptable (given the context at hand) from the no-nonsense perspective mentioned above constitute the syntactically correct occurrences of division.

### 2.3 Suppes-Ono fracterm calculus

An old idea, first described in reasonable detail by Suppes in [30] and subsequently first studied in depth by Ono in [26] is to adopt the convention that for all numbers \(x\), \(x/0 = 0\) (which I will refer to as the Suppes-Ono convention).

The Suppes-Ono convention constitutes a seemingly trivial, if not superficial, way out of the question “what is \(1/0\)” which primarily has merit in its simplicity. Although adopting the Suppes-Ono convention (that is setting \(1/0 = 0\)) fails to provide a number \(p\) so that \(0 \cdot p = 1\), which may be considered undesirable (in view of defeating the desirable “law” \(x \cdot (1/x) = 1\) for all \(x\)), the resulting arithmetic allows straightforward formalisation, which is quite helpful if one intends to implement formalised arithmetic on a computer. There is by now a significant (and steadily increasing) amount of literature about what I call Suppes-Ono fracterm
calculus (that is calculating with fractional expressions on the basis of the Suppes-Ono convention).

Although the Suppes-Ono convention has not been adopted by mathematicians (corresponding to the expectations made by Suppes in 1957) the Suppes-Ono convention has been widely adopted by logicians and informaticians in the area of automated formalisation and generation of proofs (for instance $1/0 = 0$ is adopted in the proof checking systems Coq, Isabelle, and Lean).

A systematic effort to investigate the possible role of adopting $1/0 = 0$ in mathematics is carried out by a group of Japanese workers, see e.g. [24, 27]. In [5], one may find an account of the presentation of elementary probability in the context of the Suppes-Ono convention. The latter work is based on the account of axioms for reals when adopting $1/0 = 0$ of [10]. A structure theory of fractions under this convention can be found in [11].

### 2.4 (Common) fracterm calculus

With $\perp$ an error-element is denoted. Arguably the simplest idea about division by zero is to adopt: for all $x$: $x/0 = \perp$, together with $x + \perp = \perp + x = x \cdot \perp = \perp / x = x / \perp = -\perp = \perp$. I will refer to $x/0 = \perp$ as the common meadows convention, with reference to [12, 13] where this option was first studied in detail. Although $\perp$ plays the role of an absorptive element just as Nullity ($\Phi$) in transrationals, both peripheral numbers differ significantly given that $\Phi = 0/0$ while $\Phi \neq 1/0$.

The resulting calculus of fractional expressions (common fracterm calculus, or fracterm calculus for short taking “common” as the default) enjoys so-called flattening: all but one division sign can be eliminated from any expression. Working with transrationals and Suppes-Ono fracterm calculus both do not provide fracterm flattening. From a theoretical perspective the presence of fracterm flattening constitutes a structural advantage for the conventions of common fracterm calculus which matters for theoretical work.

Including $\perp$ in a domain turns it into a so-called flat CPO (Complete Partial Order). This transformation is very common in computer science where it plays a central role in the semantics of programming languages.

The same transformation may be used for the description of partial functions with $f(a) = \perp$ representing that $f(-)$ is not defined on $a$. However, there are several alternatives to working with $\perp$ for grasping the logic of partial functions, for instance along the lines of [29], or following [21], or as in [22].

### 3 Some general observations

In this section I will look at division by zero from five perspectives: first of all I discuss why the equation $x \cdot 0 = 1$ is considered unsolvable. Secondly I will provide a brief survey of remarks from the educational literature on division by zero, and thirdly I will discuss the option that division is understood as a logical connective rather than an arithmetical operator. Subsequently I will discuss arguments for introducing fracterm as a
neologism, and finally I will confront Dedekind’s workshop with Peano’s framework.

3.1 On the equation \( x \cdot 0 = 1 \)

One may wonder how obvious it is that the equation \( x \cdot 0 = 1 \) has no solution.

1. Consider the equation \( x + 7 = 4 \). As it stands now the solvability of this equation is a matter of age: initially in school there is no such \( x \) (admittedly there is no task of solving equations either), and in due time \( 4 - 7 \) is identified as a notation for a negative number which solves the equation.

   In the history of mathematics it took centuries for the idea of negative numbers to become accepted, and since some 250 years (or even less) that it is entirely standard.

2. The equation \( x^2 = 2 \) has no rational solution. A fact known since Greek mathematics. Solving equations may come with the need to expand the domain of numbers, in this case with irrational numbers.

3. Next consider the equation \( x^2 = -1 \) or stated differently: “what is \( \sqrt{-1} \)?” Initially in school one works as if \( \sqrt{-1} \) means nothing, although most mathematicians think that \( \sqrt{-1} \) constitutes no problem at all, and that the expression has a very clear meaning, even two very clear meanings (a positive and a negative one, or rather an unsigned one and its opposite) so that a choice must be made.

   In the history of mathematics it took centuries to accept the existence of numbers which satisfy \( x^2 = -1 \), and nowadays, that is completely standard, though still somehow unknown to the public at large. It took scholars some 250 years from the first conception to the standardized use of such numbers. These numbers have been completely accepted only since about 1800.

4. There can be numbers larger than each integer number, so-called non-standard numbers. Here “can be” means that number systems including standard numbers as well as non-standard numbers are conceivable, and in fact quite well-known. The discovery of non-standard numbers is some 70 years old by now. Non-standard numbers include infinitesimals, which are above zero but below \( 1/n \) for each positive natural number \( n \) as well. Infinitesimals are inverses of infinitely large numbers.

   By construction non-standard numbers cannot solve equations for which there are no standard solutions. Thus the non-standard numbers do not allow to solve the equation \( x \cdot 0 = 1 \).

5. What I conclude from the mentioned examples is that it is not at all obvious that the equation \( x \cdot 0 = 1 \) has no solution. What can be said, though, is that number systems in which such solutions exist have not been studied in any detail thus far. Thus the unsolvability of \( x \cdot 0 = 1 \) boils down to the absence of a mathematical practice in which such solutions are possible. There is no proof whatsoever,
however, that designing and developing such a practice is impossible (unless of course it is ruled out by adopting $x \cdot 0 = 0$ as an axiom).

3.2 Comments on division by zero from the educational literature

The educational relevance of division by zero is discussed in detail in [17]. Arguments in favour of the mainstream position are surveyed in [31], the most convincing argument being referred to as the formal argument: (i) $1/b = c$ only if $b \cdot c = 1$, (ii) the zero-multiple theorem $(0 \cdot x = 0)$ is regarded a fundamental property of numbers. In [23] it is argued that $1/0$ may confuse students because it can neither denote actual infinity, nor potential infinity. This argument suggests to label the peripheral number \( \infty \) an instance of formal infinity.

In [28] a survey is given of opinions of a group of middle school teachers on division by zero, (which, it goes without saying, the authors consider impossible). The following argument was found as a teacher’s response and was considered flawed:

The math police have determined that it is against the law to divide any number by zero. This is such an important law that all calculator companies have to follow it. Here is a calculator, punch in seven divided by zero and see what happens.

In [25] the following remark suggests an awareness that division precedes its set theoretic definition.

Many of us have heard students remark,”Yes, I can see that the multiplication problem has no answer. Still, the division problem should work. Why can’t the answer be zero?” This is a natural point of view since division is an old friend, whereas the interrelationship of multiplication and division is often viewed as a newly discovered coincidence.

I sympathise with this answer. It highlights that a design decision lies at the basis of an analysis of division by zero. In [25] it is suggested that a combination of arguments provides the best option for explaining why $1/0$ is not defined.

3.3 Division: a function symbol or a logical connective?

If one writes that $x = 0 \lor x \neq 0$ there is no need or even incentive to consider the disjunction symbol $\lor$ as an arithmetical operator. It is plausible to assume that outside the scope of the equality sign(s) logic is in charge.

I will assume that it is plausible to view division as an arithmetical function, a function which may be total or partial, and that adopting the assumption that division should certainly not be understood as an arithmetical function is implausible. This paper is written under the working hypothesis that division is understood as a mathematical function.
However, admittedly one might conceive of division as constituting a part of the logic, just as negation, conjunction and disjunction are part of the logic. For instance division may be understood in some cases, that is for appropriate contexts, as a logical connective rather than as an arithmetical connective by interpreting an occurrence of $p/q$ in a context $C[-]$ by means of a transformation (where the context $C[-]$ may include/involvde logic as well as arithmetic): 

$$C[p/q] \iff q \neq 0 \land \exists x (x \cdot q = p \land C[x])$$

Even if adopting an understanding of the division symbol as a (primarily) logical connective may on the long run turn out to be the better approach to a formal understanding of division, I hold that the intuition that division is just like addition, subtraction, and multiplication one of a family of very ubiquitous arithmetical functions cannot be denied plausibility. It is with the conception of division as an arithmetical function in mind that this paper has been written, while performing a systematic investigation of division understood as a logical connective is left for another occasion.

Consider the following context $C[-]$ given by: $C[-] \equiv "- = 7"$ then $C[p/q] \equiv "p/q = 7"$. Now an interpretation of the division sign in the fracterm $p/q$ as a logical connective may yield: “$q \cdot 7 = p$”. The relevance of this apparently trivial observation comes from the fact that this particular logical interpretation of an occurrence of the division operator (symbol) demonstrates that an occurrence of the division operator symbol need not necessarily be indicative for an intention to perform an act of division (that is to compute a quotient), while such an occurrence may just as well indicate the intention to perform a multiplication in order to verify that a specific quotient has already been found.

At the logical end of the spectrum of definitions for division one may consider the use of Hilbert’s $\epsilon$ notation: $p/q = \epsilon x.(p = q \cdot x)$. For information on the $\epsilon$ notation I refer to [33].

### 3.4 Fracterm versus fraction

The term “fraction” is ambiguous. In mathematics a fraction is often a number (as in the field of fractions), and as such it cannot be equipped with a numerator and a denominator, numbers abstract from such features. In educational practice a fraction is often meant to be an expression, that is a piece of syntax which denotes a number, without being a number.

Fracterm (see [9] and [13]) denotes the same as fraction though exclusively understood as an expression, that is without the flexibility of understanding the fracterm as a number, at least not without additional explanation. Fracterm is a shorthand for what is also called: fractional expression, though in a fairly liberal sense: $1/0$ is a fracterm, whereas not all will agree that $1/0$ is a fractional expression.

The fracterm $1/0$ is qualified as not division-safe in [9] (while $(4 + 7)(5 - 1)$ is a division safe fracterm) and $1/0$ is labeled a fracpair (in addition to being a fracterm) in [13]). When we refer to $1/0$ as a fracterm then the question arises what is its meaning? And that question has no
unique answer, and in particular not the answer that as a matter of course no such meaning can exist.

If an axiomatic approach to division is contemplated, then unavoidably, one allows writing $a/b$ in advance of having determined its meaning. The freedom to do so is core to an axiomatic approach. By referring to $1/0$ as a fracterm it is implicitly assumed that fracterms (in an axiomatic approach) exist independently of the meaning assigned to such entities.

Taking a primarily syntactic view of division via the notion of a fracterm, can be transferred to addition, thus leading to the notion of a sumterm [7]. Sumterms pose additional complications with respect to fracterms, however, because sums of different lengths must be distinguished.

3.5 Dedekind’s workshop versus Peano’s framework

Richard Dedekind (1831-1916) is credited with proposing and achieving the systematic encoding of the most basic mathematical concepts in (what is now called naive) set theory. In what I call Dedekind’s workshop all mathematical notions are in essence set theoretic constructions. And each mathematician is qualified to create encodings of informal mathematical notions in set theoretic terms. A function is a set of pairs and so on. For the case of division: division is a partial function given by a set of pairs. When using a notation for it, say $a/b$ one must try to achieve best possible compliance with the underlying set theoretical understanding of division being the intended meaning of this notation.

In Peano’s framework the most basic functions (addition and multiplication) are understood by means of an axiomatic approach. Peano (1858-1932) himself was uninformed about the degree of freedom left by his axiom system when he made his design. Clarity about those matters came with Gödel’s incompleteness result in 1931.

Admittedly the construction of structures which serve as models for systems of axioms is compliant with the rules of engagement in Dedekind’s workshop. Peano’s framework does not remove the relevance of Dedekind’s workshop at all. It is not Dedekind’s workshop at the level of notions of function and relation which I wish to challenge, but merely the premature freezing of division in terms of the machinery made available in Dedekind’s workshop.

In the context of this paper it is worthwhile to notice that the flexibility of Dedekind’s approach to the definition of number systems is apparent from [3] where the use of an additional cut allows an elegant definition of transreal numbers.

4 Retrospective, prospective, and formal division

In this Section I will outline a distinction between three occurrences of division in fracterms: retrospective division, prospective division, and formal division. Formal division may alternatively be called neutral division.
Most occurrences of division in mathematics, including school arithmetic, may be classified as retrospective. Occurrences in computer programming, however, are mostly prospective, and occurrences of division in proof systems as well as in axiom systems are plausibly classified as formal. A formal occurrence of division calls in question the very idea of division, possibly leaving room for a plurality of interpretations.

The no-nonsense view alluded to in Paragraph 2.1 above is based on the idea that fracterms feature retrospective occurrences of division: \( \frac{p}{q} \) represents a number \( r \) of which it is known that \( q \neq 0 \) and \( q \cdot r = p \) which the effect that \( r \) might (retrospectively) be determined by way of division so that the notation \( \frac{p}{q} \) makes sense (after all). In other words, for each occurrence of a fracterm \( \frac{p}{q} \) some form of retrospection is required for its justification. That it be known of \( r \) that \( q \neq 0 \) may seem an implausible requirement at first sight, but said requirement makes sense as soon as \( r \) and \( q \) share some variables.

A prospective occurrence of division in a fracterm \( \frac{p}{q} \) indicates an intention to perform a division in order to determine a value \( r \) so that \( p = q \cdot r \). This intention may or may not be combined with an intention to first check that \( q \neq 0 \) and with a plan on how to proceed if it turns out that \( r = 0 \). Prospective occurrences of division occur in computer programming. Avoiding prospective divisions \( \frac{p}{q} \) to be instantiated in cases where \( q = 0 \) is both technically and conceptually a non-trivial task, which may fail for that reason, in which case one is confronted with division by zero in one way or another. Detecting in advance of a run for a program containing prospective divisions whether or not these involve division by zero is an undecidable problem.

A formal occurrence of division in a fracterm \( \frac{p}{q} \) arises when taking the context into account, and in advance of full scrutiny of the occurrence neither its qualification as prospective nor its qualification as retrospective is forced upon the reader. Formal occurrences of division may have a range of different interpretations.

Examples of formal division may arise in various settings: (i) if an axiom involves a fracterm there may be no commitment to either a retrospective or a prospective reading; (ii) if a program (or rather a program like notation) is used in order to create a logical assertion (for example as a part of a software specification) there may be no commitment to a prospective reading of division; (iii) if an unfamiliar text is investigated and a fracterm \( \frac{p}{q} \) is found in the text, it may not be known in advance which kind of occurrence is meant so that at least initially a formal interpretation may be considered most adequate.

4.1 About division in Dedekind’s workshop

The classical view on division is that division is a collection \( F_{div} \) of pairs \( ((a, b), c) \) with \( a, b, c \), taken from some number system \( K \) involving multiplication with domain \( K_s \), such that

\[
((a, b), c) \in F_{div} \iff a, b, c \in K_s \land a = b \cdot c
\]

So much is core reality about division. And then for the sake of communication about division human authors and readers use the following
convention: assuming that there is a \( c \) such that \( ((a, b), c) \in f \), write \( a/b \) for denoting \( c \).

Thus, so to say outside mathematics as a scientific subject proper (the world of sets), which is first order logic over a relation \( \in \), one may choose a notation, say \( \div \), and one may introduce the convention that given \( a \) and \( b \), if \( c \) exists such that \( ((a, b), c) \in F_{\text{div}} \) the notation \( a/b \) can be used for \( c \) (which is provably unique, provided \( 0 \neq 1 \)). In the absence of any \( c \) with \( ((a, b), c) \in F_{\text{div}} \) that notation should not be used. In the presence of parameters, with \( a, b, c \) expressions, preceding assumptions on the parameters must imply that \( b \neq 0 \).

Thus, as soon as one claims that one is involved in “doing mathematics” (at any level), one is supposed to adhere to these rules of engagement in connection with the use of notations for functions.

Unless one commits to either usage of a logic of partial functions or otherwise to the idea that all functions are total these rules of engagement stand in the way of the use of function names in axioms. Indeed, granted that set theory provides a convincing and unambiguous explanation of what a function is (a set of pairs serving as the graph of the function) it is not equally clear how set theory can explain the role of the name of a function as it appears in one or more axioms.

When it comes to multiplication, addition and subtraction, the mathematical tradition shows ample examples of the use of these function symbols in the context of axiomatisation. It is the combination of axiomatisation with partiality which seems to pose a difficulty. Indeed, that combination is not supported by first order logic. By working with a \( \bot \)-enlargement, the project of first order formalisation of partial functions is simplified, while at the same time formalisation by means of equational logic is both feasible and attractive.

A path towards a more liberal appreciation of division is brought forward by [20] where it is stated that “fraction” is not a mathematical notion. This remark gives another expression to the idea underlying the proposal to speak of fracterms explicitly. Similar ideas can be found in [32].

### 4.2 Division as a syntactic element

I will now assume that it is somehow taken for granted that in \( a/b \) the numerator \( a \) and denominator \( b \) take values in familiar number systems that is without peripheral numbers. Now the classical argument runs as follows:

(\textit{Given that we know what numbers are being considered}):

there is no \( c \) such that \( 0 \cdot c = 1 \) and therefore \( 1/0 \) is non-denoting.

There is an implicit bias in this argument, namely that \( a/b \) is only written if it is meant to denote “the unique \( c \) such that \( b \cdot c = a \)”. Precisely with this assumption semantics (that is the uniform translation of arithmetic into set theory) is given prominence at the cost of the independent status of notation.

A human author is unlikely to write \( 1/0 \) because there is no practical context in which that is useful (except in papers about division by zero).
In a similar manner a mathematical text is unlikely to contain the equation \(0 = 0\). A presence of that equation is plausible in a text on (the metatheory of) equational logic or in a formal proof ready made for automatic proof checking, but not in an ordinary mathematical text. Wittgenstein has remarked that the equality sign tends to be used between different entities only. Indeed mathematicians will only write \(a = b\) if that assertion is somehow informative, which the assertion \(a = a\) is not.

So where does the question about division by zero arise and why is it inescapable? Expressions of the form \(P/Q\) occur in computer programs and when computing, running a program, various substitution results (by substituting “concrete” values for the variables in \(P\) and in \(Q\)) of \(P\) and of \(Q\) are evaluated in preparation for evaluating \(P/Q\).

Now it may just happen that during a computation \(Q\) evaluates to 0. Avoiding such an event is not easy. Unless one restricts the freedom of (imperative programming) in a drastic manner (by requiring formal proofs for the non-occurrence of such events) the matter is just as undecidable as the famous halting problem is. (The unsolvability of the halting problem as demonstrated by Alan Turing in 1936 counts as the archetypical first significant result in theoretical computer science. In 1936 “computer” was used, however, as a reference to a human professional able to work with mechanical calculating machines.) The stochastic nature of computing as a physical process implies that actual computations (on today’s computers) may deviate from deterministic predictions so that program verification may not suffice (in principle) to prevent the occurrence of division by zero under all conceivable circumstances and with 100% certainty.

### 4.2.1 Prospective division in floating point arithmetic

So the question then becomes: what is to be done if during some computation a fraction \(P/Q\) is evaluated and \(Q\) is found to be equal to zero? Over 60 years of extensive and world-wide experience with floating point based computation over the (digitally represented) reals have now led to the following conventions (writing \(+\text{inf}\) instead of \(\infty\) as the latter notation is uncommon in computing):

- If \(P\) is positive then \(P/Q = +\text{inf}\),
- If \(P\) is negative then \(P/Q = -\text{inf}\),
- If \(P = 0\) then \(P/Q = \Phi\). Here \(\Phi\) is nullity in the datatype of transrationals, which provides a simplified model, and at the same time a simplified design, for floating point arithmetic.

These conventions make sense in a context where division is read in a prospective manner: the possibility that the denominator turns out to have value 0 must be taken into account. Doing so can be done in several ways. The above conventions have been codified in the IEEE 754 floating point standard, though I use the notation for nullity from [2].

In summary the following Proposition states circumstances which are amenable for a prospective reading of division.

**Proposition 4.1** The occurrence of an expression \(P/Q\) may express an intention to divide \(P\) by \(Q\) just as well as it may express (at the same time) the result of having (successfully) divided \(P\) by \(Q\). The intention to divide \(P\) by \(Q\) (for example in the mind of a computer programmer) may
precede becoming aware that $Q = 0$ which may in some cases stand in the way of obtaining a convincing outcome.

A programming mistake may give rise to a program fault which in turn causes the occurrence of an evaluation of $P/Q$ in circumstances where $Q$ evaluates to 0, thereby triggering an exception or a run-time failure, or the creation of a peripheral value. Needless to say that in times of ubiquitous hacking and fraud the phenomenon of intended occurrences of division by zero caused by fraudulent but seemingly innocuous occurrences of prospective division must be taken into account.

### 4.2.2 Prospective division versus intended division

Most occurrences of a fracterm $P/Q$ in computing are prospective: indications of a plan to perform actual divisions in the future, in a context where that parameters in $P$ and $Q$ have all be assigned a value. More often than not prospective division (as present in a program text) are expected to be instantiated into actual divisions many times in the run of a program, and the program is expected to run many times in the future. From the perspective of prospective division the following is uncontroversial:

**Proposition 4.2** No intended division takes the form of an occurrence of a fracterm of the form $P/0$.

The question about division by zero can be made “operational” as follows:

Given a fracterm $P/Q$, presumably involving parameters (variables). What is the value (result of evaluation) of $P/Q$ with a parameter substitution $\sigma$ to the variables in $P$ and in $Q$ for which $\sigma(Q) = 0$. (Here the substitution $\sigma$ may be unintended, that is the author of the fracterm $P/Q$ expected $P/Q$ not to be evaluated for $\sigma$).

This question matters especially in cases where it is not obvious in advance that a substitution $\sigma$ so that $\sigma(Q) = 0$ exists within the given context.

**Proposition 4.3** The precautions having to do with setting values for $1/0$ as described in Paragraph 4.2 can be entirely justified, as well as historically explained, by the phenomenon of unintended actual occurrences of division by zero which come about from instances of intentional prospective division.

One may think in terms of the notion of a descendant of a prospective division (that is a frarterm which constitutes part of a design). The presence of a fracterm in a process (for example a program) may have as descendants actual divisions among which there are instances of division by zero. Unless a program (or its compiler) performs formula manipulation, descendants of a prospective division are simply substitution instances of it.
4.2.3 Formal division for proof checking

The situation with proof checking is somewhat different. Proof checkers must determine if a proof is valid. Understanding why a proof was written the way it is, is not the issue at hand. A proof checker must be able to deal with "proofs" in which the author tries to deceive the checker. Taking for granted that in a candidate proof which involves an expression $P/Q$ the author has guaranteed beyond any possible doubt that $Q \neq 0$, defeats the very idea of proof checking. A proof checker will come across the need to check candidate proofs where $P/Q$ is used in a way which does not conform to the requirements of conventional mathematical rigour. As it turns out doing so is made easier by adopting the convention that $P/0 = 0$.

This convention, when used for proof checking can be classified as formal. Worries about adverse impact of assuming $1/0 = 0$ need not exist.

Suppose one is contemplating a proof about numbers (rational, real, complex, quaternion) which takes the form of deriving an equation $t = r$ from, say finitely many (an inessential restriction), equational assumptions collected in $E$, and suppose the proof and assertion have been written within the mainstream style of handling division by zero. Then the following can be guaranteed upon assuming $1/0 = 0$, or more generally upon assuming $x/0 = (x \cdot 1)/0 = x \cdot (1/0) = x \cdot 0 = 0$:

(i) under the "new" interpretation the assertion $t = r$ is also correct, assuming the proof (as given as input to the proof checker) was valid,

(ii) a valid proof remains valid,

(iii) a problematic proof (understood from the mainstream perspective) may become unproblematic, in which case an easy rewrite of the problematic proof will turn it into a proof which is unproblematic from the mainstream perspective as well.

**Proposition 4.4** During proof checking an occurrence of a fracterm $P/Q$ may be encountered with insufficient precautions made about $Q$ being non-zero. Provisions may be made for that situation. The (formal) reading of $x/0$ as $0$ is one of these options.

Negative proof author/designer intentions must be taken into account as well, however, and may also cause problematic use of fracterms.

4.2.4 Formal division in an axiom system

The axioms of fracterms in Table 1 constitute an example of formal use of the division sign. In advance of designing that axiom system it is unknown which models it will have.

Common fracterm calculus is the fracterm calculus of common meadows as proposed in [12] and as discussed in [13]. Table 1 lists equations for FTC following the presentation of [15]. We notice that these equations are not logically independent. Instead the design is such that: (i) the first 11 axioms can be considered as an independent module (see [15] for further information on that matter), (ii) the axioms 12-17 serve the purpose of fracterm flattening as clearly as possible, while (iii) axiom 18 is needed to obtain completeness for a meaningful class of structures.
\[(x + y) + z = x + (y + z)\]  
\[(x + y) = y + x\]  
\[x + 0 = x\]  
\[x + (-x) = 0 \cdot x\]  
\[x \cdot (y \cdot z) = (x \cdot y) \cdot z\]  
\[x \cdot y = y \cdot x\]  
\[1 \cdot x = x\]  
\[x \cdot (y + z) = (x \cdot y) + (x \cdot z)\]  
\[0 \cdot x = 0 \cdot (x \cdot x)\]  
\[x + \bot = \bot\]  
\[x = \frac{x}{1}\]  
\[-\frac{x}{y} = \frac{-x}{y}\]  
\[\frac{x}{y} \cdot \frac{u}{v} = \frac{x \cdot u}{y \cdot v}\]  
\[\frac{x}{y} + \frac{u}{v} = \frac{(x \cdot v) + (y \cdot u)}{y \cdot v}\]  
\[\frac{x}{(\frac{u}{v})} = \frac{x \cdot v \cdot v}{u \cdot v}\]  
\[\bot = \frac{1}{0}\]  
\[\frac{1}{1 + 0 \cdot x} = 1 + 0 \cdot x\]  

Table 1: FTC, fraction calculus for common meadows
4.2.5 (Formal) division as an action

One may prefer to understand a division as an action:

**Definition 4.1** A division is the act of calculating a value for a fracterm $P/Q$.

This definition is vague in various ways:

(i) it is not clear from the definition which outcomes count as values. Is $3 - 1$ acceptable as the result of calculating $8/4$. And if so, is $4/2$ an acceptable outcome of performing the division $8/4$?

(ii) Action is not a mathematical notion, although there is a plurality of theories of action in existence. It is not obvious that definition 4.1 makes sense without a notion of actor.

(iii) Speaking of action suggests a notion of state transition, again not a classical aspect of elementary mathematics.

(iv) The notion of calculation is intimately connected with the idea of preferred notations for values. So we may say that calculating a value is finding a preferred notation for the value to be calculated. Now $1/2$ may very well be the result of calculating $1/2$, but that result may also be $0.5$ depending on the preferences for notation one has embraced.

If division is meant as an action, that reading of division might well be classified as a formal reading of division, because it raises the question what type of action is involved? If the action is atomic it will not involve a test that the denominator is non-zero, and otherwise if the action is non-atomic: how does it decompose?

4.3 Summary: connecting the four story lines on division by zero with classification of division

The four views (story lines) on division by zero as outlined in Paragraph 2 above are connected with the classification of division in the following manner:

1. The no-nonsense view is based on the retrospective reading of division.

2. Taking $1/0 = +\infty$ captures the current preference for dealing with division by zero in computer programming. Division in computer programming is understood as prospective division.

3. Adopting $1/0 = 0$ provides a very simple understanding of division as formal division which is helpful for (design of and use by) proof checking systems.

The Suppes-Ono convention can in principle be used for the interpretation of prospective division, but there seems to be no example of that use of it in practice.

4. Adopting $1/0 = \bot$ provides a formal understanding of division which has the following virtues:

(i) closeness to the intuitions of the no-nonsense approach,
(ii) providing an adequate model of prospective division in most electronic calculators,
(iii) admitting a convincing axiomatisation in terms of equational logic.

The common meadows convention \(1/0 = \bot\) allows a perspective on formal division which avoids the complexity of transrationals and transreals, and which avoids the somewhat counter-intuitive simplification of the Suppes-Ono convention, while admitting fracterm flattening which matters a lot because fracterm flattening formalises an uncontested intuition from the no-nonsense approach.

5 Concluding remarks

In the concluding remarks I will first discuss the possibility of a paradigm change regarding division by zero. Subsequently I will comment on the relevance of syntax. Then I will be more precise about what constitutes the mainstream (no-nonsense) approach to division by zero. Finally I will comment on potential relevance of the plurality of perspectives on division by zero for teaching arithmetic at an elementary level.

5.1 A paradigm change on division by zero, can it happen?

A paradigm change on division by zero would occur if

(a) the prospective reading of division obtains an equal status with the currently dominant retrospective reading while leaving some limited room for a formal reading, and

(b) if a choice between these different readings must be made that choice is understood to depend on the context, while

(c) the logical intricacy of each of the different readings is acknowledged.

The mainstream position in mathematics regarding division by zero (referred to above as the no-nonsense position) is defended by a large community of individuals who seem to think that whoever challenges the mainstream position on division by zero is deeply misguided. The certainty about the mainstream position and what it has to say about division is so strong that no need is felt to be explicit about the view on division which drives the mainstream position regarding these matters.

Given today’s scientific consensus about division by zero, is a paradigm change concerning division by zero conceivable? How frequent are paradigm changes in logic and mathematics? What makes them happen? I see the following sequence of paradigm changes (without any claim to historical adequacy and completeness, just to establish that paradigm changes do occur in mathematics just as elsewhere in the sciences):

(i) Dedekind explains many concepts from mathematics in terms of (naive) set theory (around 1880).
(ii) Cantor develops set theory into an independent mathematical theory (around 1890). An axiomatic basis, for which the need is firmly established by Russell, is given in the form of Zermelo-Fraenkel (ZF) set theory (or ZFC, ZF with the axiom of choice).

(iii) Grothendieck and many others reformulate much of mathematics in terms of categories (from say 1950 to 1980).

(iv) Potentially: elementary mathematics frees itself from set theory (mainly driven by the ever increasing role of automated computing in mathematics and logic).

It is the fourth paradigm change in this list which I expect to take place in due time and which creates a context for the work in this paper. A fourth paradigm change may well be imminent, but it may work out quite differently for instance in the direction of univalent type theory which constitutes a much more ambitious and far-reaching endeavour than the path towards increased levels of by now traditional formalisation which I have in mind.

5.2 On syntax regaining 1st class status

As part of said paradigm change, I imagine loosening the grip of set theory on elementary mathematics. Understanding arithmetic as a story about set theory assigns too much importance to the semantics of the formalism, thereby creating the counter-intuitive illusion that whoever writes $7 + 5 = 12$ is in fact thinking about a fairly complicated translation $\text{pdna2ZF}("7 + 5 = 12")$ of this assertion into ZF set theory. And the same is expected of readers. Here “pdna” stands (ad hoc) for “positive decimal number arithmetic”, and “2” abbreviates “→” as a mapping. The translation from arithmetic into set theory (here named $\text{pdna2ZF}$) came about from the efforts of Dedekind and matured during the subsequent formalisation of set theory.

5.3 Principles of conventional school arithmetic

School arithmetic seems to be firmly based on a retrospective understanding of division. Some further comments are in order.

I assume that readers are comfortable with the notion of a structure, say $\mathbb{Q}(\div)$ involving zero (0), one (1), addition (+), opposite $-\cdot$, subtraction (a derived operator given by $x - y = x + (\neg y)$), multiplication ($\cdot$), and division ($\div$). $\mathbb{Q}(\div)$ is a partial algebra because division is a partial function, for no value of $x$, $x/0$ is defined (that is there is no pair of the form $((x,0), y)$ contained in the graph of the division function.

Claim 5.1 (Conventional school arithmetic on division.) Conventional school arithmetic (regarding division) constitutes an informal methodology for working with and writing about the (partial) datatype $\mathbb{Q}(\div)$ thereby making use of:

(i) the constants for decimal natural numbers,

(ii) an equality sign ($=$),

(iii) the absence of a firm commitment to the use of a two valued logic,
(iv) a tradition of avoiding making a distinction between syntax and semantics,
(v) a fairly strict imperative that expressions may only be used if the existence of a value is guaranteed by “preceding” (logically, textually, or temporally) assumptions regarding the context at hand,
(vi) an open mind regarding the form and the quality of the guarantees mentioned in (v).

It is worth mentioning that when it comes to expressions with free variables and working within the mainstream approach the notion of valid syntax becomes connected with either a notion of mathematical truth (how to know if a fraction \( P/Q \) may occur in a text), or a notion of proof (how to prove that a fraction \( P/Q \) may occur in a text). Both options impede making a clear distinction between syntax and semantics. That distinction is available for the other seemingly more complicated but in fact simpler approaches to equality in the presence of division (or of any other partial function).

**Claim 5.2** *Elementary school arithmetic defeats being formalised.*

To substantiate this claim I consider text fragments in an exposition of the form: given are some \( x_1, \ldots, x_n \in \mathbb{Z}_d \) (the decimal integer, see [7]), so which various properties may have been established already. Now we consider text fragments of the following form:

“Let the rational number \( y \) be given by \( y = (17 + u)/t(x_1, \ldots, x_n) \). We consider the quality (mainstream style adequacy) of this text fragment depending on \( t(x_1, \ldots, x_n) \). We will consider three special cases:

(i) \( t(x_1, \ldots, x_n) \equiv 275 \) in this case there is no issue because it is well-known that \( 275 \neq 0 \),
(ii) \( t(x_1, \ldots, x_4) \equiv x_1^2 + x_2^2 + x_3^2 + x_4^2 + 1 \). Again there is no problem because sums of squares in \( \mathbb{Z}_d \) will not be negative.
(iii) \( t(x_1, \ldots, x_4, y_1, \ldots, y_4, z_1, \ldots, z_4) \equiv (x_1^2 + x_2^2 + x_3^2 + x_4^2 + 1)^{587} + (y_1^2 + y_2^2 + y_3^2 + y_4^2 + 1)^{587} - (z_1^2 + z_2^2 + z_3^2 + z_4^2 + 1)^{587}.\)

Now a type checking system which easily settles this sort of matter for an arbitrary exponent is no less than an AI system able to settle Fermat’s last Theorem all by itself. Of course that particular achievement can be established by now with sufficient access to the mathematical literature, the problem having been solved. However, Hilbert’s 10th problem being undecidable in the general case there is no other way than to ask for a proof that \( t(x_1, \ldots, x_n) \) cannot be equal to 0. Formalisation won’t work: either incompleteness or inconsistency is to be expected.

### 5.4 Teaching about division by zero

Division by zero as a theme is intriguing. The main message about it may consist of the four story lines of Section 2. These matters can be taught in school at an early stage.

A survey on division by zero that may be helpful for readers who take a further interest in these matters can be found in [6]. Experiments
with the application of Suppes-Ono calculus in forensic reasoning are reported in [5]. The Suppes-Ono convention is nevertheless unattractive for teaching arithmetic at an elementary level because the clarity created by adopting $1/0 = 0$ does not outweigh (in an educational setting) the lack of intrinsic conceptual motivation for that assumption.

At the other side of the spectrum of logical complexity transrationals incorporate relevant conceptual aspects of current approaches to floating point calculation, but introducing such complexities at an elementary level of education is implausible.

I expect, however, that (common) fracterm calculus (that is the consequences of Table 1) will provide useful elements for the further development of an educational practice which goes beyond the current focus on the no-nonsense approach.

References


[19] Tiago Soares dos Reis. Transreal integral. *Transmathematica*, ISSN 2632-9212, (published 2019-06-25), [https://doi.org/10.36285/tm.v0i0.13](https://doi.org/10.36285/tm.v0i0.13), (2019).


