TYPE CLASSES FOR EFFICIENT EXACT
REAL ARITHMETIC IN COQ

ROBBERT KREBBERS AND BAS SPITTERS

Institute for Computing and Information Sciences, Radboud University Nijmegen

Abstract. Floating point operations are fast, but require continuous effort on the part of the user in order to ensure that the results are correct. This burden can be shifted away from the user by providing a library of exact analysis in which the computer handles the error estimates. Previously, we [KS11] provided a fast implementation of the exact real numbers in the Coq proof assistant. Our implementation improved on an earlier implementation by O’Connor [O’C08] by using type classes to describe an abstract specification of the underlying dense set from which the real numbers are built. In particular, we used dyadic rationals built from Coq’s machine integers to obtain a 100 times speed up of the basic operations already.

This article is a substantially expanded version of [KS11] in which the implementation is extended in the various ways. First, we implement and verify the sine and cosine function. Secondly, we create an additional implementation of the dense set based on Coq’s fast rational numbers. Thirdly, we extend the hierarchy to capture order on undecidable structures, while it was limited to decidable structures before. This hierarchy, based on type classes, allows us to share theory on the naturals, integers, rationals, dyadics, and reals in a convenient way. Finally, we obtain another dramatic speed-up by avoiding evaluation of termination proofs at runtime.

1. Introduction

There is a big gap between numerical algorithms in research papers, which typically use concepts like Hilbert spaces and fixed point theorems from functional analysis, and their actual implementation, which uses floating point1 numbers and matrix operations. This gap makes it difficult to trust the code. Similarly, this gap is undesirable in proofs of theorems (e.g. Kepler conjecture [Haa02], existence of the Lorentz attractor [Tuc02]) that rely on the execution of this code for numerical approximation. Finally, from a purely software engineering point of view, the situation is undesirable, because the gap between the (abstract

The research leading to these results has received funding from the European Union’s 7th Framework Programme under grant agreement nr. 243847 (ForMath).

1By floating points we mean numbers of the shape $n \times b^e$, where $n$ and $e$ are integers with a finite precision and $b$ is the base for scaling (typically 2, 10 or 16). The most widely used form of floating point arithmetic is the IEEE 754 standard, which is present in many hardware and software implementations.
mathematical) numerical algorithms and the (concrete floating point) implemented program makes the code difficult to maintain.

The challenge to close this gap has already been posed by Bishop in his fundamental work on constructive analysis \[\text{Bis67}\]. Bishop proposed to use constructive analysis to bridge this gap. Moreover, we can narrow this gap by using

- exact real numbers or intervals instead of floating point numbers;
- functional programming instead of imperative programming;
- dependent type theory which allows us to compute with complete metric spaces;
- a proof assistant which allows us to verify the correctness proofs;
- constructive mathematics to tightly connect mathematics with computations and to avoid computationally impossible case distinctions.

Separately, all these tools have proved itself. By going to the limits of this proven technology we should be able to come within a small constant factor of floating point computations. In this way one would obtain a tool suitable for research and education in numerical analysis that allows one to compute abstractly at the level of functional analysis, e.g. to compute fixed points of operators on Hilbert spaces. Like the development of FORTRAN and MATLAB this will require a huge amount of work. In the present paper we focus on the performance of real number computation in the COQ proof assistant.

Real numbers, being infinite objects, cannot be represented exactly in a computer. Hence, in constructive analysis \[\text{Bis67}\] one uses functions which when fed a desired precision approximate a real numbers by a rational, or a dyadic number, to within that precision.

The real numbers are the completion of the rationals. This completion construction can be organized in a monad, a familiar construct from functional programming. The completion monad provides an efficient combination of proving and computing \[\text{O'C07}\]. In this way, O’Connor \[\text{O'C08}\] implements exact real numbers and the transcendental functions on them in Coq. A number of possible improvements in this implementation were already suggested in \[\text{OS10, O'C09}\].

1. Use Coq’s new machine integers instead of the integers built from ordinary inductive data types;
2. use dyadic rationals (that are numbers of the shape \(n \ast 2^e\) for \(n, e \in \mathbb{Z}\), also known as infinitary floats) instead of ordinary rationals;
3. use approximate division to improve the implementation of power series.

Here we carry out all three optimizations. Unfortunately, changing O’Connor’s implementation to use the new machine integers was far from trivial, as he used a particular concrete representation of the rationals. To avoid this in the future, we provide an abstract specification of the dense set as approximate rationals. Finally, we obtain another dramatic speed-up by avoiding evaluation of termination proofs at runtime.

Outline. Section 2 describes some aspects of the COQ proof assistant relevant for our development. Section 3 describes metric spaces, monads, and O’Connor’s implementation of the real numbers \[\text{O'C07}\]. Section 4 extends Spitters and van der Weegen’s approach to abstract interfaces using type classes \[\text{SvdWT11}\]. Section 5 describes the theory of approximate rationals, our implementation of the real numbers, and deals with computing power series and the square root. We finish with some benchmarks in Section 6 and conclusions in Section 7. The sources of our developments can be found at \[\text{http://robbertkrebbers.nl/research/reals}\].
2. The Coq-system

The Coq proof assistant is based on the calculus of inductive constructions \cite{CH88, CP90}, a dependent type theory with (co)inductive types; see \cite{Coq08, BC04}. In true Curry-Howard fashion, it is both a pure functional programming language with an expressive type system, and a language for mathematical statements and proofs. We highlight some aspects of Coq relevant for our development.

2.1. Types and propositions. Propositions in Coq are types \cite{ML98, ML82}, which themselves have types called sorts. Coq features a distinguished sort called Prop that one may choose to use as the sort for types representing propositions. The distinguishing feature of the Prop sort is that terms of non-Prop type may not depend on the values of inhabitants of Prop types (that is, proof terms). This regime of discrimination establishes a weak form of proof irrelevance, in that changing a proof can never affect the result of value computations. On a practical level, this lets Coq safely erase all Prop components when extracting certified programs to Ocaml or Haskell. We should note however, that in practice, Coq’s extraction mechanism \cite{Let08} is still very hard to use for programs with the complexity, in terms of depth of definitions, that we are interested in \cite{CFS03, CFL06}.

2.2. Constructive indefinite description. Constructive indefinite description \cite{BC04, 14.2.3, 15.4} states that given a decidable predicate over the natural numbers, a Prop based existential can be converted into a Type based one. Its formal statement can be found in the standard library:

\[
\text{Lemma constructive_indefinite_description_nat (P : nat \rightarrow \text{Prop})}:
\forall x : \text{nat}, \{P x\} + \{\neg P x\} \rightarrow (\exists n : \text{nat}, P n) \rightarrow \{n : \text{nat} \mid P n\}
\]

Here the notation \(\{x : A \mid P x\}\) for \(P : A \rightarrow \text{Prop}\) denotes a \(\Sigma\)-type. This lemma can be seen as a form of Markov’s principle in Coq. The algorithm does a bounded search for a new witness satisfying the predicate. The witness from the Prop based existential is only used to prove termination of the search. No information flows from the Prop universe to the Type universe because the witness found for the Type based existential is independent of the witness from the Prop based one.

2.3. Equality, setoids, and rewriting. Because the Coq type theory lacks quotient types (as it endangers the decidability of type checking), one usually bases abstract structures on a setoid (‘Bishop set’): a type equipped with an equivalence relation \cite{Bis67, Hof97}. This leads to a naive set theory as described by Palmgren \cite{Pal09}. When the user attempts to substitute a given (sub)term using an equality, the system keeps track of, resolves, and combines proofs of equivalence \cite{Soz09}.

The ‘native’ notion of equality in Coq, Leibniz equality, is that of terms being convertible, naturally reified as a proposition by the inductive type family eq with single constructor eq_refl : \(\forall (T : \text{Type}) (x : T), x \equiv x\), where \(a \equiv b\) is notation for \(\text{eq} T a b\). Since convertibility is a congruence, a proof of \(a \equiv b\) lets us substitute \(b\) for \(a\) anywhere inside a term without further conditions. Our interest is in more complicated equalities, so we diverge from Coq tradition and reserve = for setoid equality. Rewriting with = does give rise to side conditions. For instance, consider formal fractions of integers as a representation of rationals. Rewriting a
subterm using such an equality is permitted only if the subterm is an argument of a function that has been proven to respect the equality. Such a function is called Proper, and that property must be proved for each function in whose arguments we wish to enable rewriting.

2.4. **Type classes.** Type classes have been a great success story in the Haskell functional programming language, as a means of organizing interfaces of abstract structures. Coq’s type classes provide a superset of their functionality, but are implemented in a different way.

In Haskell and Isabelle, type classes and their instances are second class. They are handled as specialized syntactic constructs whose semantics are given specifically by the type class apparatus. By contrast, the expressivity of dependent types and inductive families as supported in Coq, combined with the use of pre-existing technology in the system (namely proof search and implicit arguments) enable a first class type class implementation [SO08]: classes are ordinary record types (‘dictionaries’), instances are ordinary constants of these record types (registered as hints with the proof search machinery), class constraints are ordinary implicit parameters, and instance resolution is achieved by augmenting the unification algorithm to invoke ordinary proof search for implicit arguments of class type. Thus, type classes in Coq are realized by relatively minor syntactic aids that bring together existing facilities of the theory and the system into a coherent idiom, rather than by introduction of a new category of qualitatively different definitions with their own dedicated semantics.

We use the algebraic hierarchy based on type classes and its abstract specification of \( \mathbb{N}, \mathbb{Z} \) and \( \mathbb{Q} \) described in [SvdW11]. Unfortunately, we should note that we have clearly met the efficiency problems connected to the current implementation of type classes in Coq. Luckily, these efficiency problems are limited to instance resolution which is only performed at compile time. Type classes have only a very minor effect on the computation time of type checked terms due to the absence of code inlining; see Section 6 for timings.

2.5. **Virtual machine and machine integers.** Coq includes a virtual machine [GL02], `vm_compute`, based on OCaml’s virtual machine to allow efficient evaluation. Both the abstract machine and its compilation scheme have been proved correct, in Coq, with respect to the weak reduction semantics. However, we still need to extend our trusted core to a bigger kernel, as the implementation has not been formally verified.

Machine integers were also added to the Coq system [AGST10]. The usual evaluation inside Coq (`compute`) uses a special inductive type for cyclic integers, but the virtual machine uses OCaml’s machine integers. This allows for a big speed-up, for which we pay by having to trust (the virtual machine and) that OCAML treats these integers correctly. The time difference between computation with Coq’s machine integers and OCAML’s Big_int is about a factor of 20 [Spi11] on primality tests.

3. **Metric spaces**

Having completed our brief description of the Coq-system, we now turn to O’Connor’s formalization of exact real numbers [O’C07]. Traditionally, a metric space is defined as a set \( X \) with a metric function \( d : X \times X \rightarrow \mathbb{R}^+ \) satisfying certain axioms. We use a more relaxed definition of a metric space that does not require the metric be a function; see also [Ric08].
The metric is represented via a (respectful) ball relation $B : Q_+ \to X \to X \to \text{Prop}$ satisfying:

- $\text{msp\_refl} : \forall x \varepsilon, B \varepsilon x x$
- $\text{msp\_sym} : \forall x y \varepsilon, B \varepsilon x y \to B \varepsilon y x$
- $\text{msp\_triangle} : \forall x y z \varepsilon_1 \varepsilon_2, B \varepsilon_1 x y \to B \varepsilon_2 y z \to B \varepsilon_1 \varepsilon_2 x z$
- $\text{msp\_closed} : \forall x y \varepsilon, (\forall \delta, B \varepsilon + \delta x y) \to B \varepsilon x y$
- $\text{msp\_eq} : \forall x y, (\forall \varepsilon, B \varepsilon x y) \to x = y$

The ball relation $B \varepsilon x y$ expresses that the points $x$ and $y$ are within $\varepsilon$ of each other. We call this a ball relationship because the partially applied relation $B^X x : X \to \text{Prop}$ is a predicate that represents the closed ball of radius $\varepsilon$ around the point $x$. For example, the ball relation on $Q$ is $B^Q x y := |x - y| \leq \varepsilon$.

A metric space $X$ is a prelength space if:

- $\forall a b \varepsilon \delta_1 \delta_2, \varepsilon < \delta_1 + \delta_2 \to B \varepsilon a b \to \exists c, B \delta_1 a c \land B \delta_2 c b$.

This property states that if two points $a$ and $b$ within $\varepsilon$ of each other, then there exists curve of length $d(a, b)$ in the completion of $X$ that connects $a$ and $b$. The metric space $Q$ is a prelength space.

We will introduce the completion of a metric space as a monad. In order to do this we will first introduce monads.

3.1. Monads. Moggi [Mog89] recognized that many non-standard forms of computation may be modeled by monads\(^2\). Wadler [Wad92] popularized their use in functional programming. Monads are now an established tool to structure computation with side-effects. For instance, programs with input $X$ and output $Y$ which have access to a mutable state $S$ can be modeled as functions of type $X \times S \to Y \times S$, or equivalently $X \to (Y \times S)^S$. The type constructor $M Y := (Y \times S)^S$ is an example of a monad. Similarly, partial functions may be modeled by maps $X \to Y_\perp$, where $Y_\perp := Y + ()$ is a monad.

The formal definition of a (strong) monad is a triple $(M, \text{return}, \text{bind})$ consisting of a type constructor $M$ and two functions:

- $\text{return} : X \to MX$
- $\text{bind} : (X \to MY) \to MX \to MY$

We will denote $\text{return} x$ as $\hat{x}$, and $\text{bind} f$ as $\hat{f}$. These two operations must satisfy:

- $\hat{\text{bind} \text{return} a} = a$
- $\hat{f \circ g} a = \text{bind}(\hat{f} \circ \hat{g}) a$

\(^2\)In category theory one would speak about the Kleisli category of a (strong) monad.
3.2. **Completion monad.** The completion of a metric space $X$ is defined by:

$$
\mathcal{C}X := \{ f : \mathbb{Q}^+ \to X \mid \forall \varepsilon_1 \varepsilon_2, B_{\varepsilon_1 + \varepsilon_2} (f \varepsilon_1) (f \varepsilon_2) \}.
$$

Given metric spaces $X$ and $Y$, a function $f : X \to Y$ is **uniformly continuous** with modulus $\mu_f : \mathbb{Q}^+ \to \mathbb{Q}^+$ if:

$$
\forall x_1 x_2, B_{\mu_f \varepsilon} x_1 x_2 \to B_{\varepsilon} (f x_1) (f x_2).
$$

Completion is a monad on the category of metric spaces with uniformly continuous functions. The function $\text{return} : X \to \mathcal{C}X$ defined by $\lambda x \varepsilon, x$ is the embedding of a metric space in its completion. Moreover, a uniformly continuous function $f : \mathcal{C}X \to \mathcal{C}Y$ with modulus $\mu_f$ can be lifted to operate on complete metric spaces as $\text{bind} f : \mathcal{C}X \to \mathcal{C}Y$ defined by $\lambda x \varepsilon, f (x (\mu_f \varepsilon)) \varepsilon$. A restriction to prelength spaces is essential for this efficient definition of $\text{bind}$.

One advantage of this approach is that it helps us to work with simple representations. Let $\mathbb{R} := \mathcal{C}\mathbb{Q}$. Then to specify a function from $\mathbb{R} \to \mathbb{R}$, we define a uniformly continuous function $f : \mathbb{Q} \to \mathbb{R}$, and obtain $\tilde{f} : \mathbb{R} \to \mathbb{R}$ as the required function. Hence, the completion monad allows us to do in a structured way what was already folklore in constructive mathematics: to work with simple, often decidable, approximations to continuous objects.

4. **Abstract interfaces using type classes**

An important part of this work is to further develop the algebraic hierarchy based on type classes by Spitters and van der Weegen [SvdW11]. Especially, we extend their hierarchy with constructive fields, order theory and interfaces for mathematical operations, such as shift and power, common in programming languages. This layer of abstraction makes both proof engineering and programming more flexible: it avoids duplication of code, it introduces a canonical way to refer to operations and properties, both by names and notations, and it allows us to easily swap different implementations of number representations and their operations. First we will briefly recap the design decisions made in [SvdW11].

Algebraic structures are expressed in terms of a number of carrier sets, a number of relations and operations, and a number of laws that the operations satisfy. One way of describing such a structure is by a **bundled representation**: one uses a dependently typed record that contains the carrier, operations and properties. For example a semigroup can be represented as follows. (The fields $\text{sg\_car}$ and $\text{sg\_proper}$ support our explicit handling of naive set theory in type theory.)

```plaintext
Record SemiGroup : Type := {
  sg\_car : Setoid ;
  sg\_op : sg\_car \to sg\_car \to sg\_car ;
  sg\_proper : Proper (\equiv) \equiv (\equiv) sg\_op ;
  sg\_ass : \forall x y z, sg\_op x (sg\_op y z) = sg\_op (sg\_op x y) z )
}
```

However, this approach has some serious limitations, the most important one being a lack of support for **sharing** components. For example, suppose we group together two **CommutativeMonoids** in order to create a **SemiRing**. Now awkward hacks are necessary to establish equality between the carriers. A second problem is that if we stack up these records to represent higher structures the projection paths become increasingly long.

Historically these problems have been an acceptable trade-off because **unbundled representations**, in which the carrier and operations are parameterized, introduce even more problems.
Record SemiGroup {A} (e : A → A → Prop) (sg.op : A → A → A) : Prop :=
{ sg_proper : Proper (e ⇒ e ⇒ e) sg.op ;
  sg_ass : ∀ x y z, e (sg.op x (sg.op y z)) (sg.op (sg.op x y) z) }

There is nothing to bind notation to, no structure inference, and declaring and passing requires too much manual bookkeeping. Spitters and van der Weegen have proposed a use of Coq’s new type class machinery that resolves many of the problems of unbundled representations. Our current experiment confirms that this is a viable approach.

An alternative solution is provided by packed classes [GGMR09] which use an alternative, and older, implementation of a semblance of type classes: canonical structures; see also Section 7. Yet another approach would be to use modules. However, as these are not first-class, we would be unable to define, e.g. homomorphisms between algebraic structures.

An operational type class is defined for each operation and relation.

Class Equiv A := equiv: relation A.
Infix "=" := equiv: type_scope.
Class RingPlus A := ring_plus: A → A → A.
Infix "+" := ring_plus.

Now an algebraic structure is just a type class living in Prop that is parametrized by its carrier, relations and operations. This class contains all laws that the operations should satisfy. Since the operations are unbundled we can easily support sharing. For example let us consider the SemiRing interface.

Class SemiRing A {e : Equiv A} {plus: RingPlus A}
{ mult: RingMult A} { zero: RingZero A}
{ one: RingOne A} : Prop :=
{ semiring_mult_monoid : @CommutativeMonoid A e mult one ;
  semiring_plus_monoid : @CommutativeMonoid A e plus zero ;
  semiring_distr : Distribute (.*.) (+) ;
  semiring_left_absorb : LeftAbsorb (.*.) 0 }. 

Without type classes it would be a burden to manually carry around the carrier, relations and operations. However, because these parameters are just type class instances, the type class machinery will perform that job for us. For example, Lemma example’ {SemiRing R} x : 1 * x = x + 0.

The backtick instructs Coq to automatically insert implicit declarations, namely e plus
mult zero one. It also lets us omit a name for the SemiRing R parameter itself. All of these parameters will be given automatically generated names that we will never refer to. Furthermore, instance resolution will automatically find instances of the operational type classes for the written notations. Thus the above is really:

Lemma example {R e plus mult zero one} {P : @SemiRing R e plus mult zero one} x : @equiv R e
(@ring_mult R mult (@ring_one R one) x)
(@ring_plus R plus x (@ring_zero R zero)).

The syntax :: in the definition of SemiRing declares certain fields as substructures. This syntax should not be confused with the similar syntax for coercions in records (e.g. in the bundled representation of a SemiGroup on page 6). In current versions of Coq, inference of substructures is based on backward reasoning. In this example that means, each time a CommutativeMonoid R instance is needed, instance search may try to find a SemiRing R instance. This style of instance search presents some problems, as the following example illustrates.

Class Setoid_Morphism {A B} {Ae: Equiv A} {Be: Equiv B} (f: A → B) :=

Each time we have to establish Setoid R, instance search might try to find a Setoid_Morphism to or from R. Since this quickly results in a serious blow-up, we omit the $\to$ declaration. Support for forward reasoning would solve this problem. If we would be in a context in which we know something to be a Setoid_Morphism, then forward reasoning automatically infers that the source and target are Setoids. Ongoing work by Matthieu Sozeau should make this style of instance search available in Coq.

Proving that an actual structure is an instance of the SemiRing interface is straightforward. First we define instances of the operational type classes.

```
Instance nat_equiv: Equiv nat := eq.
Instance nat_plus: RingPlus nat := plus.
Instance nat_0: RingZero nat := 0%nat.
Instance nat_1: RingOne nat := 1%nat.
Instance nat_mult: RingMult nat := mult.
```

Here we see that instances are just ordinary constants of the class types. However, we use the `Instance` keyword instead of `Definition` to let the type class machinery register the instance.

Now, to prove that the Peano naturals are in fact a semiring, we just write:

```
Instance: SemiRing nat.
Proof. ... Qed.
```

This approach to interfaces proved useful to formalize a standard algebraic hierarchy. Combined with category theory and universal algebra, $\mathbb{N}$ and $\mathbb{Z}$ are represented as interfaces specifying an initial semiring and initial ring [SvdW11].

```
Class NaturalsToSemiRing (A : Type) :=
  naturals_to_semiring: \forall \{\text{SemiRing } B\}, SemiRing_Morphism (naturals_to_semiring A B);
  naturals_initial: Initial (semirings.object A).
```

These abstract interfaces for the naturals and integers make it easier to change the concrete representation in the future. No such simple specification for $\mathbb{Q}$ seems to exists, so we choose to specify it as the field of fractions of $\mathbb{Z}$. More precisely, $\mathbb{Q}$ is specified as a field containing $\mathbb{Z}$ that moreover can be embedded into the field of fractions of $\mathbb{Z}$.

```
Inductive Frac R \{\text{e : Equiv } R\} \{\text{zero : RingZero } R\} : Type :=
  frac {\text{num : } R ; \text{den : } R ; \text{den_nonzero : } \text{den} \neq 0}.
Class RationalsToFrac (A : Type) :=
  rationals_to_frac : \forall \{\text{Integers } B\}, A \to Frac B.
Class Rationals A \{\text{e plus mult zero one opp inv}\} \{\text{U : !RationalsToFrac A}\} : Prop :=
  rationals_field: \forall \{\text{DecField } A \text{ e plus mult zero one opp inv}\};
  rationals_frac: \forall \{\text{Integers } Z\}, Injective (rationals_to_frac A Z);  
  rationals_embed_ints: \forall \{\text{Integers } Z\}, Injective (integers_to_ring Z A).
```

4.1. **Constructive fields and apartness.** In constructive mathematics, the common notion of inequality as the negation of equality is often too weak because a proof of a negation lacks computational content. For example, in order to define the reciprocal of $x \in \mathbb{R}$, one
needs a witness $\varepsilon \in \mathbb{Q}_+$ that $|x| \geq \varepsilon$. Such a witness cannot be extracted from a proof of $x \neq 0$. To solve this problem, one uses a setoid equipped with an apartness (irreflexive, asymmetric and co-transitive) relation describing inequality $\text{[TvD88]}$.

The algebraic hierarchy in the CoRN library $\text{[CFGW04]}$ has been built on top of such setoids. Unfortunately, this hierarchy is quite ‘heavy’in practice. First, for structures with decidable equality, the negation of equality is the only tight apartness. Hence, when working with decidable structures, an apartness relation is unnecessary. Secondly, CoRN uses an informative (that is, $\text{Type}$ based) apartness relation to facilitate extraction of witnesses. However, CoQ’s present implementation of setoid rewriting does not support rewriting over relations in $\text{Type}$. So, it does not allow us to replace equations in expressions involving CoRN’s informative apartness and thus many proofs involve a lot of manual labor.

To remedy these issues we propose an alternative solution. We use a non-informative (that is, $\text{Prop}$-based) apartness relation to enable setoid rewriting and include it just in the parts of the algebraic hierarchy where we actually need it. The latter keeps our interfaces clean and easy to use and should combine the best of two worlds. Type classes are of great help to reduce bookkeeping and clutter in proofs.

Although using a non-informative apartness relation enables setoid rewriting, it disables extraction of witnesses. Fortunately, in case of the reals, a witness can be obtained inefficiently by bounded linear search (see Section 2.2 and 5.1). We think our approach is a reasonable trade-off since the amount of reasoning exceeds the potential use of apartness in computation. In case we need a witness for efficient computation, we just have to specify it explicitly. This approach of specifying witnesses explicitly was already preferred by O’Connor $\text{[O’C08]}$, even when an informative apartness was available.

Our interface for a setoid with apartness (henceforth $\text{StrongSetoid}$) is as follows.

\begin{verbatim}
Class Apart A := apart: relation A.
Infix "< >" := apart (at level 70, no associativity) : type_scope.

Class StrongSetoid A {e: Equiv A} {ap : Apart A} : Prop := 
  {\text{strong_setoid_irreflexive} : Irreflexive (<>)} ;
  {\text{strong_setoid_symmetric} : Symmetric (<>) ;}
  {\text{strong_setoid_cotrans} : CoTransitive (>)} ;
  {\text{tight_apart} : \forall x y, \neg x < > y \iff x = y}.

This interface is equipped with a tight equality. We prove that each $\text{StrongSetoid}$ is a $\text{Setoid}$. For decidable structures, we define the following class to describe that the apartness relation is the negation of equality.

\begin{verbatim}
Class TrivialApart R {Equiv R} {ap : Apart R} := trivial_apart : \forall x y, x <> y \iff x \neq y.
\end{verbatim}

Given a setoid with decidable equality we can easily extend it to a $\text{StrongSetoid}$.

\begin{verbatim}
Instance default_apart `{\text{Equiv A}} : Apart A | 20 := (\neq).
Instance default_apart_trivial `{\text{Equiv A}} : TrivialApart A (ap:=default_apart).
Lemma dec_strong_setoid `{\text{Setoid A}} `{\text{Apart A}}
  `{\text{TrivialApart A}} `{\forall x y, \text{Decision}(x = y)} : StrongSetoid A.
\end{verbatim}

Unfortunately, the type class mechanism is unable to detect simple loops. Hence we define the above as an ordinary $\text{Lemma}$ instead of an $\text{Instance}$. This trick prevents CoQ from using it in instance search and therefore avoids endless derivations of the form $\text{StrongSetoid A, Setoid A, StrongSetoid A, \ldots}$.
For ordinary setoids we want functions to be **Proper**, which means that they respect equality. For setoids with apartness we need a stronger property, **strong extensionality**.

**Class** StrongSetoid Morphism \{ A B : Type \} \{ Ae : Equiv A \} \{ Aap : Apart A \}

\{ Be : Equiv B \} \{ Bap : Apart B \} \{ f : A \to B \} := \{ 

\text{strong_setoidmor}_a : \text{StrongSetoid} A ; \\
\text{strong_setoidmor}_b : \text{StrongSetoid} B ; \\
\text{strong_extensionality} : \forall x y, f x \gg f y \to x \gg y \}.

We prove that for each StrongSetoid Morphism \( f \) we have \( \text{Proper} ((=) \implies (=)) f \). The only structures for which we actually need apartness are implementations of the real numbers, hence we only base the Field class on top of a StrongSetoid instead of the complete algebraic hierarchy.

Our class for fields is as follows. (The PropHolds class is explained in the next subsection.)

**Class** Field A \{ e plus mult zero one opp \} \{ ap : Apart A \} \{ mult_inv : MultInv A \} := Prop := \{ 

\text{field_ring} : @Ring A e plus mult zero one opp ; \\
\text{field_strongsetoid} : \text{StrongSetoid} A ; \\
\text{field_plus_ext} : \text{StrongSetoid BinaryMorphism} (+) ; \\
\text{field_mult_ext} : \text{StrongSetoid BinaryMorphism} (\ast.) ; \\
\text{field_nontrivial} : \text{PropHolds} (1 \gg 0) ; \\
\text{mult_inv_proper} : \text{Setoid Morphism} (/) ; \\
\text{mult_inverse} : \forall x, 'x \div x = 1 \}.

We do not include strong extensionality of the inverse and the reciprocal because it can be derived. For convenience, we define a class for fields with decidable equality whose reciprocal function is total. This class integrates nicely with Coq’s rational numbers \( \mathbb{Q} \) and \( \mathbb{Z} \), and the field tactic. This total reciprocal function should satisfy \( /0 = 0 \), so properties as \( f((/x) = (/f(x)), (/x) = x \) and \( /x * /y = (/x * y) \) hold without any additional premises.

A diagram of our complete algebraic hierarchy is displayed in Figure 1.

4.2. **Order theory.** Previous Coq libraries for ordered algebraic structures turn out to be too limited to abstract from \( \mathbb{N}, \mathbb{Z}, \mathbb{Q} \) and \( \mathbb{R} \) and their various implementations. The formalization of ordered fields in the CoRN library [CFGW04] restricts to a very specific part of the algebraic hierarchy (namely fields). Letouzey’s NUMBERS library, which is included in recent versions of Coq trunk, only considers \( \mathbb{N} \) and \( \mathbb{Z} \). The SSREFLECT library presently restricts to decidable structures with Leibniz equality. Moreover, even mathematically, the most convenient abstraction is not entirely clear. Lešnık [Les10] provides a smooth order theoretic characterization of these structures as so-called streaks. We, however, prefer our theory below as it avoids partial functions.

In this work we present a library that captures the notion of order on a variety of structures, including structures with undecidable equality. One of the building blocks of our hierarchy is a pseudo order, which is the constructive variant of a total order.

**Class** PseudoOrder \{ e : Equiv A \} \{ ap : Apart A \} \{ so : Lt A \} := Prop := \{ 

\text{pseudo_order_setoid} : \text{StrongSetoid} A ; \\
\text{pseudo_order_antisym} : \forall x y, \neg (x < y \land y < x) ; \\
\text{pseudo_order_cotrans} : \text{CoTransitive} (<) ; \\
\text{apart_iff_totalLt} : \forall x y, x \gg y \leftrightarrow x < y \lor y < x \}.

In case equality is decidable, this interface is rather awkward to work with. Therefore we present ways to go back and forth between the usual classical notions and their constructive variants. For example, we use the type class machinery to infer the classical trichotomy property.
Instance It_trichotomy '(PseudoOrder A) '('TrivialApart A) '('∀ x y, Decision (x = y)) : Trichotomy (<).

Also, we can go the other way around. If we have a StrictSetoidOrder (an ordinary strict order built upon a setoid) satisfying the trichotomy property, we obtain a pseudo order.

Lemma dec_strict_pseudo_order ' (StrictSetoidOrder A) '(Apart A) '-' (TrivialApart A) '('∀ x y, Decision (x = y)) '-' (Trichotomy (<)): PseudoOrder (<).

In order to avoid loops, we define the above as an ordinary Lemma instead of an Instance. Next, one could extend a pseudo order to the standard notion of a (pseudo) ring order.

Class PseudoRingOrder ' (Equiv A) '(Apart A) '(RingPlus A) '(RingMult A) '(RingZero A) (Alt : Lt A) := {
    pseudo_ringorder_spo := PseudoOrder Alt ;
    pseudo_ringorder_mult_ext := StrongSetoid_BinaryMorphism (.* ) ;
    pseudo_ringorder_plus := ∀ z, StrictlyOrderPreserving (z + ) ;
    pseudo_ringorder_mult := ∀ x y, 0 < x → 0 < y → 0 < x * y }.

However, we would like to apply our order library to implementations of the naturals, which are merely semirings, too. Therefore, we strengthen the above with a partial subtraction function (living in Prop, because we never use it for computations) and require addition to be order reflecting. We call this, apparently new notion, a PseudoSemiRingOrder.

Class PseudoSemiRingOrder ' (Equiv A) '(Apart A) '(RingPlus A) '
   '(RingMult A) '(RingZero A) (Alt : Lt A) := {
    pseudo_srorder_strict := PseudoOrder Alt ;
    pseudo_srorder_partial_minus := ∀ x y, ¬ y < x → ∃ z, y = x + z ;
    pseudo_srorder_plus := ∀ z, StrictOrderEmbedding (z + ) ;
    pseudo_srorder_mult_ext := StrongSetoid_BinaryMorphism (.* ) ;
pseudo_sorder_pos_mult_compat : \forall y y', 0 < y \rightarrow 0 < y' \rightarrow 0 < x + y'.

Instead of including the PseudoRingOrder class in our development, we include a lemma to construct a PseudoSemiRingOrder from a ring satisfying the PseudoRingOrder axioms.

Given a pseudo (semiring) order, one could define the non-strict order \( x \leq y \) in terms of the strict order, namely as \( \neg y < x \). However, this is quite inconvenient in practice, because we also want to talk about a priori different non-strict orders such as those defined in the standard library. Hence we introduce the following class.

Class \( \text{FullPseudoSemiRingOrder} \) \(
\{ \text{Equiv A}, \text{Apart A}, \text{RingPlus A}, \text{RingMult A}, \text{RingZero A} \} \)
\(
\{ \text{Le : Le A} \}, \text{Alt : Lt A} \} := \{
\text{full_pseudo_sorder_pso} : : \text{PseudoSemiRingOrder Alt} ;
\text{full_pseudo_sorder_le_iff_not_lt_flip} : \forall x y, x \leq y \leftrightarrow \neg y < x.
\)

A diagram of our complete order hierarchy is displayed in Figure 1.

Our theory on abstract orders avoids duplication of theorems and proofs. For example, the following lemmas apply to \( \mathbb{N}, \mathbb{Z}, \mathbb{Q} \) and the dyadics, because all of these structures form a FullPseudoSemiRingOrder.

Lemma \( \text{plus_compat} x_1 y_1 x_2 y_2 : x_1 \leq y_1 \rightarrow x_2 \leq y_2 \rightarrow x_1 + x_2 \leq y_1 + y_2 \).

Lemma \( \text{lt_1_2} : 1 < 2 \).

Lemma \( \text{square_nonneg} x : 0 \leq x \times x \).

To allow us to refer by canonical names to common properties, we introduce classes like those shown below:

Class \( \text{OrderPreserving} \) \( \{ A B \} \{ \text{Ae: Equiv A}, \text{Ale: Le A}, \text{Be: Equiv B}, \text{Ble: Le B} \} \) \( \{ f : A \rightarrow B \} := \{
\text{order_preserving_morphism} := \text{OrderMorphism} ;
\text{order_preserving : } (x \leq y \rightarrow f x \leq f y).
\)

Class \( \text{OrderReflecting} \) \( \{ A B \} \{ \text{Ae: Equiv A}, \text{Ale: Le A}, \text{Be: Equiv B}, \text{Ble: Le B} \} \) \( \{ f : A \rightarrow B \} := \{
\text{order_preserving_back_morphism} := \text{OrderMorphism} ;
\text{order_preserving_back : } (f x \leq f y \rightarrow x \leq y).
\)

Here, an OrderMorphism is just the factoring out of the common parts of both classes; namely that \( f \) and \( \leq \) respect equality. For the case of multiplication these properties have additional premises, for example:

Global Instance: \( \forall (z : \mathbb{R}), \text{PropHolds (0 < z)} \rightarrow \text{OrderPreserving (z *.)} \).

We introduce the PropHolds class to let the type class machinery prove these properties automatically. For example consider:

Lemma \( \text{example (n : N)} (x y : \mathbb{R}) : x \leq y \rightarrow (2 \times n + 2) \times x \leq (2 \times n + 2) \times y \).

Proof. intros. now apply (order_preserving \( (2 \times n + 2) \times \)). Qed.

In order to use order_preserving, we need a proof of PropHolds \( (0 < 2 \times n + 2) \). Type class resolution is able to prove this in a fully automated way because we have the following instances:

Instance: PropHolds \( (0 < 2) \);

Instance: \( \forall x y : \mathbb{R}, \text{PropHolds (0 < x)} \rightarrow \text{PropHolds (0 < y)} \rightarrow \text{PropHolds (0 < x + y)} \);

Instance: \( \forall (n : N) (x : \mathbb{R}), \text{PropHolds (0 < x)} \rightarrow \text{PropHolds (0 < x * n)} \).

For arbitrary instances of \( \mathbb{N}, \mathbb{Z}, \mathbb{Q} \) it is easy to define an order satisfying these interfaces:

Instance nat_le \( \{ \text{Naturals N} \} : \text{Le N} | 10 := \lambda x y, \exists z, y = x + z \).

Instance nat_lt \( \{ \text{Naturals N} \} : \text{Lt N} | 10 := \lambda x y, x \leq y \land x \neq y \).

However, often we encounter an a priori different order on a structure, most likely an order defined in Coq’s standard library (like \( \text{Nle} \) and \( \text{Nil} \) on \( \mathbb{N} \)). Therefore we prove that a FullPseudoSemiRingOrder uniquely specifies the order on \( \mathbb{N}, \mathbb{Z} \) and \( \mathbb{Q} \). For example:
Context ‘Naturals N {Naturals N2} {f : N → N2} ‘{SemiRing_Morphism f} ‘{Apart N} ‘{TrivialApart N} ‘{FullPseudoSemiRingOrder (A:=N) Nle Nlt} ‘{Apart N2} ‘{TrivialApart N2} ‘{FullPseudoSemiRingOrder (A:=N2) N2le N2lt}.

Global Instance: OrderEmbedding f.

Unfortunately Coq has no support to have an argument be ‘inferred if possible, generalized otherwise’; see [SvdW11]. When declaring a parameter of FullPseudoSemiRingOrder, one is often in a context where most of its components are already available. Usually, only the parameters Le, Lt and Apart have to be introduced. The current workaround in these cases involves providing names for components that are then never referred to, which is a bit awkward. In the above it would much nicer to write:

Context ‘Naturals N {Naturals N2} {f : N → N2} ‘{SemiRing_Morphism f} ‘{TrivialApart N} ‘{FullPseudoSemiRingOrder N} ‘{TrivialApart N2} ‘{FullPseudoSemiRingOrder N2}.

Global Instance: OrderEmbedding f.

4.3. Basic operations. The operation nat.pow is most commonly, but inefficiently, defined as repeated multiplication and the operation shiftl is defined as repeated duplication. Instead we specify the desired behavior of these operations. This approach allows for different implementations for different number representations and avoids definitions and proofs becoming implementation dependent.

We introduce interfaces that specify the behavior of the operations abs, shiftl, nat.pow and int.pow. Again there are various ways of specifying these interfaces: with Σ-types, bundled or unbundled. In general, Σ-types are convenient for functions whose specification is easy, for example:

Class Abs A ‘{Equiv A} ‘{Le A} ‘{RingZero A} ‘{GroupInv A}
   := abs_sig: ∀ (x : A), { y : A | (0 ≤ x → y = x) ∧ (x ≤ 0 → y = −x)}.

Definition abs ‘{Abs A} := λ x : A, ‘(abs_sig x).

However, for more complex operations, such as shiftl, such an interface is different from the usual mathematical specification because we cannot quantify over all possible input values. Now there are two ways: a bundled or an unbundled interface. Since these interfaces are not used for hierarchies the disadvantages of the former do not apply. Let us first describe the former approach.

Class ShiftL A B ‘{Equiv A} ‘{Equiv B} ‘{RingOne A} ‘{RingPlus A} ‘{RingMult B} ‘{RingOne B} ‘{RingPlus B} :=
   { shiftl : A → B → A ;
     shiftl_proper : Proper ((=)⇒(=)⇒(=)) shiftl ;
     shiftl_0 : RightIdentity shiftl 0 ;
     shiftl_S : ∀ x n, shiftl x (1 + n) = 2 * shiftl x n }.

Infix “≪” := shiftl (at level 33, left associativity).

Although this interface seems reasonable, it does not work well in Coq. The simpl tactic which is used to simplify a goal will unfold occurrences of shiftl to their underlying definition (for example in case of BigN, the expression x ≪ n becomes BigN.shiftl x n). This is rather inconvenient because Coq will then be unable to use lemmas concerning ≪ for rewriting. This problem is caused because shiftl is a projection of a record, which is in fact an ι-redex (reduction of pattern-matching over a constructed term) that will be unfolded by simpl. Currently there seems to be no way to adjust the behavior of simpl to remove this inconvenience. A similar problem was already observed in Ssreflect [GMT08].
Instead we use an unbundled interface, which has a lot in common with our interfaces for algebraic structures. Now \( \text{shiftl} \) no longer contains an \( \iota \)-redex.

\[
\text{Class } \text{ShiftL } A B := \text{shiftl} : A \rightarrow B \rightarrow A.
\]

\[
\text{Infix } \preceq := \text{shiftl} \text{ (at level 33, left associativity).}
\]

\[
\text{Class } \text{ShiftLSpec } A B (sl : \text{ShiftL } A B) \{ \text{Equiv } A \} \{ \text{Equiv } B \} \{ \text{RingOne } A \} \{ \text{RingPlus } A \} \{ \text{RingMult } A \} \{ \text{RingZero } B \} \{ \text{RingOne } B \} \{ \text{RingPlus } B \} := \{
\text{shiftl} \text{proper} : \text{Proper } ((=) \implies (=) \implies (=)) \preceq ;
\text{shiftl}0 := \text{RightIdentity } (\preceq) 0 ;
\text{shiftl}S := \forall x n, x \preceq (1 + n) = 2 \ast x \preceq n \}.
\]

We do not specify \( \text{shiftl} \) as \( \text{shiftl} x n = x \ast 2 \ast n \) since on the dyadics we cannot take a negative power while we can shift by a negative integer. Since theory on shifting with exponents in \( \mathbb{N} \) and \( \mathbb{Z} \) is similar we want to avoid duplication of theorems and proofs. To this end we introduce a class describing the bi-induction principle.

\[
\text{Class } \text{Biinduction } R \{ \text{Equiv } R \} \{ \text{RingZero } R \} \{ \text{RingOne } R \} \{ \text{RingPlus } R \} \{ \text{RingMult } R \} \{ \text{RingZero } R \} \{ \text{RingOne } R \} \{ \text{RingPlus } R \} \{ \text{RingMult } R \} : \text{Prop} := \text{biinduction } (P : R \rightarrow \text{Prop}) \{ \text{!Proper } ((=)\implies iff) P \} : P 0 \rightarrow (\forall n, P n \iff P (1 + n)) \rightarrow \forall n, P n.
\]

4.4. Decision procedures. The Decision type class collects types with a decidable equality \cite{SvdW11}.

\[
\text{Class } \text{Decision } P := \text{decide} : \text{sumbool } P (\neg P).
\]

Using this type class we can declare a parameter \( \forall x y, \text{Decision } (x \leq y) \) to describe a decider for \( \leq \) and say \text{decide} \( (x \leq y) \) to decide whether \( x \leq y \) or not. This type class allows us to easily define additional deciders, like the one for the strict order. We have to be careful however. Consider the order on the dyadics.

\[
\text{Global Instance } \text{dy preceded} : \text{Le Dyadic} := \lambda (x y : \text{Dyadic}), \text{ZtoQ } (\text{mant } x) + 2^{\ast} (\text{expo } x) \leq \text{ZtoQ } (\text{mant } y) + 2^{\ast} (\text{expo } y)
\]

Now, \text{decide} \( (x \leq y) \) is actually \text{@decide Dyadic } (x \leq y) \text{dyadic_dec}, where \text{dyadic_dec} is the computational conclusion of the decision. Due to eager evaluation, and the absence of dead code removal, the second argument, \( x \leq y \), is also evaluated. Evaluation of this argument results in a conversion of \( x \) and \( y \) into \( \mathbb{Q} \), as described above. But since this argument is just a proposition it is later thrown away. We avoid this problem introducing a \( \lambda \)-abstraction.

\[
\text{Definition } \text{decide rel } (R : \text{relation } A) \{ \text{dec} : \forall x y, \text{Decision } (R x y) \}
\]
\[
(x y : A) : \text{Decision } (R x y) := \text{dec } x y.
\]

We can now define:

\[
\text{Context } \{ \text{FullPseudoOrder } A \} \{ \text{TrivialApart } A \} \{ \forall x y, \text{Decision } (x \leq y) \}.
\]

\[
\text{Global Program Instance } \text{lt dec} : \forall x y, \text{Decision } (x < y) | \exists := \lambda x y,
\]
\[
\text{match } \text{decide rel } (\leq) x y \text{ with}
\]
\[
| \text{left } E \Rightarrow \text{right }_.
\]
\[
| \text{right } E \Rightarrow \text{left }_.
\]
\[
\text{end.}
\]

Notice that this problem also occurs if we had defined the following bool-based variant.
4.5. **Explicit type casts.** The `Cast` type class collects (explicit) type casts.

```coq
Class Cast A B := cast: A -> B.
Implicit Arguments cast [[Cast]].
Notation " x " := (cast _ _ x) (at level 20).
Instance: Params (@cast) 3.
```

This definition allows us to refer to a cast from `x : A` to `B` by using an apostrophe, or writing `cast A B x`. An example of an instance of this class is:

```coq
Instance NonNegInject: Cast (R^+) R := @proj1 sig R _.
```

Here, `R^+` is the non-negative cone of an ordered ring `R`. Contrary to Coq’s built-in coercion mechanism, our type casts are explicit instead of implicit and type classes are used to register them. Our approach has some advantages:

1. By using type classes to register casts, we are allowed to parametrize classes with casts. An example is the `AppRationals` class, as defined in Section 5.
2. Implicit coercions often introduce ambiguity. Since our approach allows us to refer to casts by a (canonical) name, e.g. `cast B C (cast A B x)`, we can avoid this ambiguity.
3. Casts can be put in partially applied position, e.g. `order_preserving (cast Z Q)`.

Coq’s coercion mechanism does not allow us to define a coercion from `R^+` to `R` nor a coercion from a ring to its polynomial ring. More generally, it does not allow most forms of parametrized coercions nor non-uniform coercions. If Coq would allow parametrized coercions like `NonNegInject`, it would have to avoid an infinite loop: to type check `x : R`, we try to type check `x : (R^+)`, but to do so we try `x : (R^+)^`, ... We suffer from such loops if we compose our `Cast` classes automatically as well. Hence we refrain from adding:

```coq
Instance cast_comp_base `f : Cast A B` : ComposedCast A B := f.
Instance cast_comp_step `f : Cast B C ` `g : ComposedCast A B` : ComposedCast A C := \lambda x, f (g x).
```

Matita [ASCTZ07] allows parametrized coercions and avoids the loop by not applying coercions recursively, but instead building a well-chosen set of set of composite coercions [Tas08]. Non-uniform coercions [ST11] are available in Matita. They are implemented using unification hints, a feature similar to type classes.

5. The real numbers

To make our implementation of the reals independent of the underlying dense set, we provide an abstract specification of approximate rationals inspired by the notion of approximate fields which is used in the Haskell implementation of the exact reals by Bauer and Kavler [BK09]. In particular, we provide an implementation of this interface by dyadics based on Coq’s machine integers.

Our interface describes an ordered ring containing `Z` that is dense in `Q`. Here `Z` are the binary integers from Coq’s standard library, and `Q` are the rationals based on these binary integers. We do not parametrize by arbitrary integer and rational implementations because they are hardly used for computation.
Also, for efficient computation, this interface contains the operations: approximate division, normalization, an embedding of \( \mathbb{Z} \), absolute value, power by \( \mathbb{N} \), shift by \( \mathbb{Z} \), and decision procedures for both equality and order.

\[
\text{Class } \text{AppDiv } AQ := \text{app_div} : AQ \rightarrow AQ \rightarrow \mathbb{Z} \rightarrow AQ.
\]

\[
\text{Class } \text{AppApprox } AQ := \text{app_approx} : AQ \rightarrow \mathbb{Z} \rightarrow AQ.
\]

\[
\text{Class } \text{AppRationals } AQ \{ \text{e plus mult zero one inv} \} \{ \text{Apart } AQ \} \{ \text{Le } AQ \} \{ \text{Lt } AQ \}
\]

\[
\{ \text{AQtoQ} : \text{Cast } AQ \rightarrow \mathbb{Q} \text{ as MetricSpace} \} \{ \text{AppInverse } AQtoQ \} \{ \text{ZtoAQ} : \text{Cast } \mathbb{Z} \rightarrow AQ \}
\]

\[
\{ \forall x y : AQ, \text{Decision } (x = y) \} \{ \forall x y : AQ, \text{Decision } (x \leq y) \} : \text{Prop} :=
\]

\[
\text{aq} \text{ring} := @\text{Ring } AQ \ e \text{ plus mult zero one inv} ; \\
\text{aq trivial apart} := \text{TrivialApart } AQ ; \\
\text{aq order embed} := \text{OrderEmbedding } AQtoQ ; \\
\text{aq dense embedding} := \text{DenseEmbedding } AQtoQ ; \\
\text{aq div} := \forall x y k, \text{ball } 2^k (\text{app div x y k}) (x / y) ; \\
\text{aq compress} := \forall x k, \text{ball } 2^k (\text{app approx x k}) (x) ; \\
\text{aq shift} := \text{ShiftSpec } AQ Z (\leq) ; \\
\text{aq nat pow} := \text{NatPowSpec } AQ N (\cdot) ; \\
\text{aq ints mor} := \text{SemiRing Morphism } ZtoAQ .
\]

Following O’Connor [O’C07], we define the real numbers as the completion of the approximate rationals. To create functions on the real numbers, we use the monadic operations \( \text{bind} \) or \( \text{map} \). This approach is convenient because equality and inequality are decidable on the approximate rationals, whereas it is not on the real numbers. For binary functions, e.g. addition and multiplication, we use the \( \text{map2} \) function, as described in [O’C07].

O’Connor [O’C07] keeps the size of the rational numbers small to avoid efficiency problems. He introduces a function \( \text{approx } x \epsilon \) that yields the ‘simplest’ rational number between \( x - \epsilon \) and \( x + \epsilon \). We modify the \( \text{approx} \) function slightly: \( \text{app approx x k} \) yields an arbitrary element between \( x - 2^k \) and \( x + 2^k \). Using this function we define the compress operation on the real numbers:

\[
\text{compress} := \text{bind} \ (\lambda x \epsilon, \text{app approx x (Qdlog2 } \epsilon)) \text{ such that } \text{compress } x = x .
\]

In Section 5.4 we will explain our choice of using a power of 2 to specify the precision of \( \text{app div} \) and \( \text{app approx} \).

5.1. Order and apartness. Following [BB85] [O’C09], we define non-negativity and the order on the real numbers as follows.

\[
\text{NonNeg } x := \forall \epsilon : \mathbb{Q}_+, -\epsilon \leq x \epsilon \\
x \leq y := \text{NonNeg } (y - x)
\]

Bishop and Bridges [BB85] define positivity as the dual of non-negativity: \( \exists \epsilon : \mathbb{Q}_+, \epsilon < x \epsilon . \) O’Connor [O’C09] defines positivity and the strict order differently so as to avoid a potentially expensive computation, namely \( x - x \epsilon , \) to obtain a witness between 0 and \( x \).

\[
\text{Pos } x := \{ \epsilon : \mathbb{Q}_+ | \epsilon \leq x \} \\
x <_T y := \text{Pos } (y - x)
\]

We use the \( T \) subscript to emphasize that the relation lives in \( \text{Type} \). Next, we define \( x \gg_T y := x <_T y \lor y <_T x \). Extraction of a witness \( \epsilon \in (0,x) \) from \( \text{Pos } x \) allows us to define the reciprocal function of type \( \forall x : \mathbb{R}, 0 \gg_T x \rightarrow \mathbb{R} \).
In order to use our type class based hierarchy we need a strict order and apartness relation in Prop. We need this restriction because Coq’s present implementation of setoid rewriting does not allow rewriting in Type-based relations (see Section 4.1). Our definition is similar to Bishop and Bridges’, but defined in such a way that we can easily prove a correspondence with O’Connor’s.

\[ x < y := \exists n : \mathbb{N}, \ 1 \leq -n \leq (y - x) \ (1 \leq -(n - 1)) \]
\[ x \succ y := x < y \lor y < x \]

Using constructive indefinite description (see Section 2.2), it is an easy job to prove that we indeed have \( x < y \iff x <_T y \) and \( x \succ y \iff x >_T y \). Similar to O’Connor [O’C09], we implement a tactic that automatically proves strict inequalities. The tactic terminates iff the inequality holds and has quite some similarities with our use of linear search to obtain \( x <_T y \) from \( x < y \).

5.2. Implementation using the dyadics. The dyadic rationals are numbers of the shape \( n \cdot 2^e \) for \( n, e \in \mathbb{Z} \). In order to remain independent of an integers implementation, we abstract over it. For our eventual implementation of the approximate rationals we use Coq’s machine integers, \( \text{bigZ} \). Now given an arbitrary integer implementation \( \text{Int} \) it is straightforward to define the dyadics. Here we will just show the ring operations.

Notation \( "x \ | \ p" := (\text{exist}_x \ p) \) (at level 20).
Record Dyadic := dyadic { mant : Int ; expo : Int }.
Infix "$" := dyadic (at level 80).
Global Instance dy_inject: Cast Int Dyadic := \( \lambda x, x \ | \ 0 \).
Global Instance dy_opp: GroupInv Dyadic := \( \lambda x, \text{mant} x \ | \ \text{expo} x \).
Global Instance dy_mult: RingMult Dyadic := \( \lambda x y, \text{mant} x \ | \ \text{mant} y \ | \ \text{expo} x \ | \ \text{expo} y \).
Global Instance dy_0: RingZero Dyadic := cast Int Dyadic 0.
Global Instance dy_1: RingOne Dyadic := cast Int Dyadic 1.
Global Program Instance dy_plus: RingPlus Dyadic := \( \lambda x y, \)
  if decide_rel (\( \leq \)) (\text{expo} x) (\text{expo} y)
  then \( \text{mant} x \ | \ \text{mant} y \ | \ (\text{expo} y \ | \ \text{expo} x) \ | \ \text{min} (\text{expo} x) (\text{expo} y) \)
  else \( \text{mant} x \ | \ (\text{expo} x \ | \ \text{expo} y) \ | \ \text{mant} y \ | \ \text{min} (\text{expo} x) (\text{expo} y) \).

In this code \text{shiftl} has type \( \text{Int} \rightarrow \text{Int}^+ \rightarrow \text{Int} \), where \( \text{Int}^+ \) is a \( \Sigma \)-type describing the non-negative cone of \( \text{Int} \). Therefore, in the definition of dy_plus we have to equip \( \text{exp} \) with a proof that it is in fact non-negative.

5.3. Implementation using the rationals. Our development contains additional implementations of the \text{AppRationals} class using Coq’s old rational numbers \( \mathbb{Q} \) and the new rational numbers \( \text{bigQ} \) (which are built from the machine integers \( \text{bigZ} \)). Although creating these implementations is uninteresting from a performance point of view, it confirms that it is trivial to change the underlying dense set from which our real numbers are built.

To implement the \text{app.approx} function in an efficient manner, we use shifts on the underlying integers. Furthermore, to keep the size of the results of the division operation small, we incorporate the \text{app.approx} function.

Instance \( \text{bigQ.div} \) : \( \text{AppDiv bigQ} := \lambda x y, \text{app.approx} (x / y) \).
5.4. **Power series.** Elementary transcendental functions as exp, sin, ln and atan can be defined by their power series. If the coefficients of a power series are alternating, decreasing and have limit 0, then we obtain a fast converging sequence with an easy termination proof. For $-1 \leq x \leq 0$,

$$\exp x = \sum_{i=0}^{\infty} \frac{x^i}{i!}$$

is of this form. To approximate exp $x$ with error $\varepsilon$ we take the partial sum until $\frac{x^i}{i!} \leq \varepsilon$.

In order to implement this efficiently we use a stream representing the series and define a function that sums the required number of elements. For example, the series $1, a, a^2, a^3, \ldots$ is defined by the following stream.

**CoFixpoint** powers\_help (c : A) : Stream A := Cons c (powers\_help (c \times a)).

**Definition** powers : Stream A := powers\_help 1.

Streams in Coq, like lists in Haskell, are lazy. So, in the example the multiplications are accumulated.

Since Coq only allows structural recursion (and guarded co-recursion) it requires some work to convince Coq that our algorithm terminates. Intuitively, one would describe the limit as an upperbound of the required number of elements using the \texttt{Exists} predicate.

**Inductive** Exists A (P : Stream A \rightarrow Prop) (x : Stream) : Prop :=

| Here : P x \rightarrow Exists P x |
| Further : Exists P (tl x) \rightarrow Exists P x.

This approach leads to performance problems. The upperbound, encoded in unary format, may become very large while generally only a few terms are necessary. Due to \texttt{vm\_compute}'s eager evaluation scheme, this unary number will be computed before summing the series. Instead O'Connor [O'C09] uses \texttt{LazyExists}.

**Inductive** LazyExists A (P : Stream A \rightarrow Prop) (x : Stream A) : Prop :=

| LazyHere : P x \rightarrow LazyExists P x |
| LazyFurther : (unit \rightarrow LazyExists P (tl x)) \rightarrow LazyExists P x.

Unfortunately, our experiments showed that the above still yields too much overhead due unnecessary to reduction of proofs. To remedy this issue we introduce the following function where \texttt{Str\_nth\_tl n s} takes the \texttt{n}-th tail of the stream \texttt{s}.

**Fixpoint** LazyExists\_inc 'P (n : nat) s : LazyExists P (Str\_nth\_tl n s) \rightarrow LazyExists P s :=

match n return LazyExists P (Str\_nth\_tl n s) \rightarrow LazyExists P s with

| O \Rightarrow λ x, x |
| S n \Rightarrow λ ex, LazyFurther (λ _, LazyExists\_inc n (tl s) ex)

end.

This function adds \texttt{n} additional \texttt{LazyFurther} constructors. When instantiated with a big enough \texttt{n}, computation will suffer from the implementation limits of Coq (e.g. a stack overflow) or runs out of memory, before it ever refers to the actual proof. Using \texttt{LazyExists\_inc} we are able to compute on average twice the amount of decimals as we did before on examples such as the ones in Table 2.

O’Connor’s \texttt{InfiniteAlternatingSum s} returns the real number represented by the infinite alternating sum over \texttt{s}, where the stream \texttt{s} is decreasing, non-negative and has limit 0. We extend this in two ways. First, we generalize various notions to abstract structures.
Secondly, as we do not have exact division on approximate rationals, we extend the algorithm to work with approximate division. The latter requires changing $\text{InfiniteAlternatingSum}$ to $\text{InfiniteAlternatingSum\ nd}$ which computes the infinite alternating sum of the stream $\lambda_i, \frac{\varepsilon}{2^k}$.

This allows us to postpone divisions. Also, we have to determine both the length of the partial sum and the required precision of the divisions. To do so we find a $k$ such that:

$$B_{\frac{\varepsilon}{2}}(\text{app\_div}\ n_k\ d_k\ (\log\frac{\varepsilon}{2^k}) + \frac{\varepsilon}{2^k}) 0.$$  \hspace{1cm} (5.1)

Now $k$ is the length of the partial sum, and $\frac{\varepsilon}{2^k}$ is the required precision of division. Using O’Connor’s results we have verified that these values are correct and such a $k$ indeed exists for a decreasing, non-negative stream with limit 0.

As noted in Section 5, we have specified the precision of division in powers of 2 instead of using a rational value. This allows us to replace (5.1) with:

$$B_{\frac{\varepsilon}{2}}(\text{app\_div}\ n_k\ d_k\ (\log\varepsilon - (k + 1)) + 1 \ll (\log\varepsilon - (k + 1))) 0.$$  \hspace{1cm} (5.2)

Here $k$ is the length of the partial sum, and $2^l$, where $l = \log\varepsilon - (k + 1)$, is the required precision of division. This variant can be implemented without any arithmetic on the rationals and is thus much more efficient.

This method gives us a fast way to compute the infinite alternating sum, in practice, only a few extra terms have to be computed and due to the approximate division the auxiliary results are kept as small as possible.

Similarly, using this method to compute infinite alternating sums, we use the following series to implement $\tan x$ and $\sin x$ for $x \in [-1, 1]$.

$$\tan x = \sum_{i=0}^{\infty} (-1)^i * x^{2i+1} \ (2i+1)!$$

$$\sin x = \sum_{i=0}^{\infty} (-1)^i * \frac{x^{2i+1}}{2i+1}$$

We extend these functions to their complete domain by repeatedly applying the following formulas $[O’C09]$.

$$\exp x = (\exp (x \ll 1))^2$$  \hspace{1cm} (5.2)

$$\exp x = \frac{1}{\exp (-x)}$$  \hspace{1cm} (5.3)

$$\sin x = 3 * \sin \frac{x}{3} - 4 * \left(\sin \frac{x}{3}\right)^3$$  \hspace{1cm} (5.4)

$$\tan x = -\tan (-x)$$  \hspace{1cm} (5.5)

$$\tan x = \frac{\pi}{2} - \tan \frac{1}{x} \ \text{for} \ 0 < x$$  \hspace{1cm} (5.6)

$$\tan x = \frac{\pi}{4} - \tan \left(\frac{x - 1}{x + 1}\right) \ \text{for} \ 0 < x$$  \hspace{1cm} (5.7)

Since we do not have exact division on the approximate rationals, we parameterize infinite sums by two streams in Equation 5.4, 5.6 and 5.7.

The series described in this section converge faster for arguments closer to 0. We use Equation 5.2 and 5.4 repeatedly to reduce the input to a value $|x| \in [0, 2^k]$. For $50 \leq k$, this yields nearly always major performance improvements, for higher precisions setting it
to 75 \leq k \) yields even better results. Unfortunately, we are unaware of a similar trick for \( \text{atan} \). We define \( \pi \) in terms of \( \text{atan} \) using the following Machin-like formula.

\[
\pi := 176 \times \text{atan} \frac{1}{57} + 28 \times \text{atan} \frac{1}{239} - 48 \times \text{atan} \frac{1}{682} + 96 \times \text{atan} \frac{1}{12943}
\]

Again, here we notice the purpose of parameterizing infinite sums by two streams. We define \( \cos \) in terms of \( \sin \).

\[
\cos x = 1 - 2 \times \left( \sin \frac{x}{2} \right)^2
\]

O’Connor \cite{O’C07, O’C09} subtracts multiples of \( 2\pi \) to reduce the arguments of \( \sin \) and \( \cos \). In our tests this did not lead to performance improvements because our implementation of \( \pi \) turned out to be slower than the performed range reductions.

5.5. **Square root.** We use Wolfram’s algorithm \cite[p.913]{Wol02} for computing the square root. Its complexity is linear, in fact it provides a new binary digit in each step.

\begin{verbatim}
Context '(Pa : 1 \leq a \leq 4).
Fixpoint AQroot_loop (n : nat) : AQ \times AQ :=
  match n with
  | O \Rightarrow (a, 0)
  | S n \Rightarrow
    let (r, s) := AQroot_loop n in
    if decide_rel (\leq) (s + 1) r
    then ((r - (s + 1)) \ll (2:Z), (s + 2) \ll (1:Z))
    else (r \ll (2:Z), s \ll (1:Z))
  end.

Let us write \((r_n, s_n)\) for the \( n \)-th pair of approximations. We prove the following facts by induction:

\[
s_n^2 + 4r_n = 4^{n+1}a \quad \text{(5.8)}
\]

\[
r_n \leq 2s_n + 4 \quad \text{(5.9)}
\]

\[
2^n s_n \leq s_{n+m} \leq 2^n (s_n + 4) - 4 \quad \text{(5.10)}
\]

\[
r_n \leq 2^{3+n} \quad \text{(5.11)}
\]

By 5.8 \((2^{-(n+1)}s_n)^2 + 2^{-2n}r_n = a\). By 5.11 \(2^{-2n}r_n\) converges to 0 as \(n\) tends to \(\infty\). Therefore, by 5.10 \(2^{-(n+1)}s_n\) is a Cauchy sequence which furthermore converges to the root of \(a\).

Next, we extend the square root to its entire domain by repeatedly applying the following formula:

\[
\sqrt{x} = 2 \times \sqrt{\frac{x}{4}}
\]

O’Connor’s Coq implementation \cite{O’C08} includes the much faster Newton iteration, whose complexity is logarithmic in the number of decimals. The function to iterate is:

\begin{verbatim}
Definition f (r : Q) : Q := r / 2 + a / (2 * r).
\end{verbatim}

Because of the absence of exact division on our approximate rationals we cannot implement this function directly. However, we can implement it on our real numbers. As the above definition does not use sharing, we have to define this function on the reals by first defining:

\begin{verbatim}
Definition f (r : AQ) (c : Qpos) : AQ := (r + approx_div (Qdlog2 c) a) \ll (-1).
\end{verbatim}
and then showing that it gives rise to a continuous function $f : \mathbb{Q} \to \mathbb{R}$ which we finally lift to a function $\text{bind} : \mathbb{R} \to \mathbb{R}$ on the reals. In this way we take care of sharing, division and intermediate use of the $\text{approx}$ function (see Section 5) all in one go. We hope the future correctness proof to be quite smooth, since we work with exact real numbers. We have implemented this in Haskell, which performs really well.

6. Benchmarks

The first step in this research was to create a Haskell prototype based on O’Connor’s implementation of the real numbers in Haskell [O’C07]. The second step was to implement and verify this prototype in Coq. Our Coq development contains verified versions of: the field operations, exponentiation by a natural, computation of power series, $\exp$, $\text{atan}$, $\sin$, $\cos$, $\pi$ and the square root.

In this section we present some benchmarks, taken from the ‘Many Digits’ friendly competition [NW09], comparing the old and the new implementation, both in Haskell and Coq. All benchmarks have been carried out on an Intel Core Quad 2.4 GHz with 8GB of memory running Debian GNU/Linux with kernel 2.6.38. The sources of our developments can be found at [http://robbertkrebbers.nl/research/reals](http://robbertkrebbers.nl/research/reals).

<table>
<thead>
<tr>
<th>Expression</th>
<th>Decimals</th>
<th>Old</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01 $\sin(\sin(\sin 1))$</td>
<td>5.000</td>
<td>25s</td>
<td>2.3s</td>
</tr>
<tr>
<td>P02 $\sqrt{\pi}$</td>
<td>5.000</td>
<td>3.3s</td>
<td>1.7s</td>
</tr>
<tr>
<td>P03 $\sin e$</td>
<td>5.000</td>
<td>13s</td>
<td>1.2s</td>
</tr>
<tr>
<td>P04 $\exp(\pi \times \sqrt{163})$</td>
<td>5.000</td>
<td>22s</td>
<td>2.0s</td>
</tr>
<tr>
<td>P05 $\exp(\exp e)$</td>
<td>5.000</td>
<td>43s</td>
<td>2.6s</td>
</tr>
<tr>
<td>P06 $\log(1 + \log(1 + \log(1 + \log(1 + \pi))))$</td>
<td>500</td>
<td>107s</td>
<td>2.5s</td>
</tr>
<tr>
<td>P07 $\exp 1000$</td>
<td>20.000</td>
<td>1.1s</td>
<td>0.7s</td>
</tr>
<tr>
<td>P08 $\cos(10^{50})$</td>
<td>20.000</td>
<td>6.7s</td>
<td>1.4s</td>
</tr>
<tr>
<td>P09 $\sin(3 \times \log_{\sqrt{163}}(500))$</td>
<td>5.000</td>
<td>33s</td>
<td>16s</td>
</tr>
<tr>
<td>P11 $\tan e + \text{atan} e + \tanh e + \text{atanh} \frac{1}{e}$</td>
<td>500</td>
<td>41s</td>
<td>3.2s</td>
</tr>
<tr>
<td>P12 $\text{asin} \frac{1}{e} + \text{cosh} e + \text{asinh} e$</td>
<td>500</td>
<td>99s</td>
<td>3.2s</td>
</tr>
</tbody>
</table>

Table 1. Haskell, compiled with ghc version 6.12.1, using -O2.

<table>
<thead>
<tr>
<th>Expression</th>
<th>Decimals</th>
<th>Old</th>
<th>New</th>
<th>Decimals</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>P01 $\sin(\sin(\sin 1))$</td>
<td>25</td>
<td>46s</td>
<td>0.6s</td>
<td>500</td>
<td>3.8s</td>
</tr>
<tr>
<td>P02 $\sqrt{\pi}$</td>
<td>25</td>
<td>0.3s</td>
<td>0.03s</td>
<td>500</td>
<td>6.8s</td>
</tr>
<tr>
<td>P03 $\sin e$</td>
<td>25</td>
<td>36s</td>
<td>0.1s</td>
<td>500</td>
<td>1.9s</td>
</tr>
<tr>
<td>P04 $\exp(\pi \times \sqrt{163})$</td>
<td>10</td>
<td>214s</td>
<td>0.1s</td>
<td>500</td>
<td>3.7s</td>
</tr>
<tr>
<td>P05 $\exp(\exp e)$</td>
<td>10</td>
<td>36s</td>
<td>0.2s</td>
<td>500</td>
<td>3.2s</td>
</tr>
<tr>
<td>P07 $\exp 1000$</td>
<td>10</td>
<td>2662s</td>
<td>1.0s</td>
<td>20.000</td>
<td>4.9s</td>
</tr>
<tr>
<td>P08 $\cos(10^{50})$</td>
<td>25</td>
<td>11s</td>
<td>0.3s</td>
<td>20.000</td>
<td>12s</td>
</tr>
</tbody>
</table>

Table 2. Coq trunk, revision 14023.
Table 1 shows some benchmarks in Haskell with compiler optimizations enabled (~O2) and Table 2 compares our Coq implementation with O’Connor’s. More extensive benchmarking shows that our Haskell implementation generally benefits from a 15 times speed up while the speed up in Coq is generally more than a 100 times for small examples already. This difference between the comparison of the Haskell and the Coq implementation is explained by the fact that O’Connor’s Haskell implementation already uses rational numbers built from fast integers and incorporates various optimizations, while his Coq implementation does not. The last column of Table 2 indicates that our new implementation is able to compute an order of magnitude more decimals in the same amount of time.

We also compared the new reals built from Coq’s fast rationals (Section 5.3) and our dyadic rationals (Section 5.2). For exp, sin and cos we obtain quite similar results due to the our range reductions to reduce the length of the power series. In case of the square root, the dyadics rationals are much faster because Wolfram iteration is designed for an efficient shift. It is interesting to notice that π and atan benefit the least from our improvements, as we are unaware of range reductions to reduce the length of the series.

We conclude this section with a comparison between the performance of Wolfram’s algorithm in Coq and Haskell. The Haskell prototype (without compiler optimizations) is quite fast, computing 10,000 iterations (giving 3,010 decimals) of √2 takes 0.2s. In Coq it takes 7.4s using type classes and 7.2s without type classes. Here we exclude the time spent on type class resolution. Thus type classes cause only a 3% performance penalty on computations.

Unfortunately, the Coq implementation is slow compared to Haskell. Laurent Théry suggested that this is due to the representation of the fast integers, which uses a tree with a fixed depth and when the size of the integer becomes too big uses a less optimal representation. Increasing the size of the tree representation and avoiding an inefficiency in the implementation of shifts reduces this time to 5.4s.

7. Conclusions and Related work

We have greatly improved the performance of real number computation in Coq using Coq’s new machine integers. We produced highly structured and abstract code using type classes with no apparent performance penalty. Moreover, Coq’s notation mechanism combined with unicode characters gives nicely readable statements and proofs. Type classes were a great help in our work. However, the current implementation of instance resolution is still experimental and at times too slow (at compile time).

Canonical structures provide an alternative, and partially complementary, implementation of type classes [GZND11]. By choice, canonical structures restrict to deterministic proof search, this makes them more efficient, but also somewhat more intricate to use. The use of canonical structures by the Ssreflect team [GGMR09] makes it plausible that with some effort we could have used canonical structures for our work instead. However, the Ssreflect-library is currently not suited for setoids which are crucial to us. The integration of unification hints [ARCT09] into Coq may allow a tighter integration of type classes and canonical structures.

We needed to adapt our correctness proofs to prevent the virtual machine from eagerly evaluating them. Lazy evaluation for Prop would have allowed us to use the original proofs. Moreover, setoid rewriting over relations in Type would have made our work much easier.
The experimental native compute performs evaluation by compilation to native OCAML code. This approach uses the OCAML compiler available and is interesting for heavy compilation. Our first experiments indicate a 10 times speed up with Wolfram iteration. Unfortunately, native compute does not work with COQ trunk yet, so we were unable to test it with our implementation of the reals.

The Flocq project [BM11] formalizes infinitary floating-points in COQ. It provides a library of theorems on multi-radix multi-precision arithmetic and supports efficient numerical computations inside COQ. However, the current library is still too limited for our purposes, but in the future it should be possible to show that they form an instance of our approximate rationals. This may allow us to gain some speed by taking advantage of fine grained algorithms instead of our more straightforward ones.

The encoding of real numbers as streams of ‘bits’ is potentially interesting. However, currently there is a big difference in performance. The computation of 37 decimals of the square root of 1/2 by Newton iteration [JP09], using the framework described in [Ber07], took 12s. This should be compared with our use of the Wolfram iteration, which gives only linear convergence, but with which we nevertheless obtain 3,000 decimals in a similar time. On the other hand, the efficiency of \( \pi \) in their framework is comparable with ours. Berger [Ber09], too, uses co-induction for exact real computation.

The present work is part of a larger program to use constructive mathematics based on type as a programming language for exact analysis. This should culminate in a numerical ODE-solver. As part of this goal, we plan to build a type class interface for metric spaces.

Cohen and Mahboubi [CM11] provide a formalization of quantifier elimination for the theory of decidable real closed fields. Together with the formalization of cylindrical algebraic decomposition [Mah07] this should lead to an efficient decision procedure. Such a procedure would have been useful in our current formalization effort.

Acknowledgements. We thank Eelis van der Weegen for many discussions and Pierre Letouzey and Matthieu Sozeau for closing some of our bug reports.

References


