The Dedekind Reals in Abstract Stone Duality

Andrej Bauer and Paul Taylor

3 June 2009

Abstract

Abstract Stone Duality (ASD) is a direct axiomatisation of general topology, in contrast to the traditional and all other contemporary approaches, which rely on a prior notion of discrete set, type or object of a topos.

ASD reconciles mathematical and computational viewpoints, providing an inherently computable calculus that does not sacrifice key properties of real analysis such as compactness of the closed interval. Previous theories of recursive analysis failed to do this because they were based on points; ASD succeeds because, like locale theory and formal topology, it is founded on the algebra of open subspaces.

ASD is presented as a lambda-calculus, of which we provide a self-contained summary, as the foundational background has been investigated in earlier work.

The core of the paper constructs the real line using two-sided Dedekind cuts. We show that the closed interval is compact and overt, where these concepts are defined using quantifiers. Further topics, such as the Intermediate Value Theorem, are presented in a separate paper that builds on this one.

The interval domain plays an important foundational role. However, we see intervals as generalised Dedekind cuts, which *underly* the construction of the real line, not as sets or pairs of real numbers.

We make a thorough study of arithmetic, in which our operations are more complicated than Moore's, because we work constructively, and we also consider back-to-front (Kaucher) intervals.

Finally, we compare ASD with other systems of constructive and computable topology and analysis.

AMS classification: 03F60 06E15 18C20 26E40 54D45 65G40.

Keywords: Dedekind real numbers, abstract Stone duality, Heine–Borel theorem, compact topological space, constructive analysis, interval analysis, domain theory, lambda calculus, exact real arithmetic, Kaucher arithmetic.

Contents

To appear in *Mathematical Structures in Computer Science*, **19** (2009).

1 Introduction

The title of Richard Dedekind's paper [Ded72] leads with the word "Stetigkeit", officially translated as "continuity". Irrational numbers get second billing — his construction gives us access to them only *via* continuity, and he stresses the importance of geometrical intuition. In other words, the real line is not a naked *set* of Dedekind cuts, dressed by later mathematicians in an outfit of so-called "open" subsets, but has a *topology* right from its conception. Dedekind cited square roots as an example of the way in which we use continuity to enrich the rational numbers, but even the rationals — being defined by division — presuppose an inequality relation that we shall come to regard as topology.

Although a great deal of mathematics has since been built over Dedekind's *construction*, we still have no *definition* of the real numbers that is widely accepted across different foundational settings. This is in contrast to the Dedekind–Peano–Lawvere definition of the *natural* numbers, which has been applied within logical systems that are much weaker than the classical one, or differ substantially from it in other ways. This is a very unfortunate state of affairs at a time when a debate has at last begun amongst a number of separate disciplines that all call themselves "constructive", ours being one of them. We need a definition so that we can agree on what we're talking about.

So, at the risk of seeming presumptuous in the face of such a venerable object, let us first write down our own opinion of what it is that we are trying to construct. In this paper we shall use R for the object under construction, and $\mathbb{R}$ for the real line in classical or other forms of analysis.

Definition 1.1 An object R is a ***Dedekind real line*** if
(a) it is an overt space (Theorem 9.2);
(b) it is Hausdorff, with an inequality or apartness relation, $\neq$ (Theorem 9.3);
(c) the closed interval $[0, 1]$ is compact (Theorem 10.7);
(d) R has a total order, that is, $(x \neq y) \Leftrightarrow (x < y) \vee (y < x)$ (Theorem 9.3);
(e) it is Dedekind-complete, in a sense in which the two halves of a cut are open (Theorem 9.6);
(f) it is a field, where x^{-1} is defined iff $x \neq 0$, in the sense of (d) (Theorem 13.4);
(g) and Archimedean (Theorem 13.6), that is, for $x, y : R$,

$$y > 0 \;\Rightarrow\; \exists n{:}\mathbb{Z}.\; y(n-1) < x < y(n+1).$$

Our axioms are all true of the classical real line. Indeed, with the exception of the new concept of overtness, they are all *headline* properties in traditional analysis — just as induction had been formulated two centuries before Dedekind and Peano encapsulated it in their axioms, and used two millennia before that. These are not just peculiar order-theoretic facts that happen to lend themselves to some interesting construction.

However, some of our constructive colleagues have not adopted certain of these axioms. In particular, many formalist accounts and machine implementations use Cauchy sequences instead of Dedekind cuts. You already know our opinion on this question from the title of this paper: familiar examples such as Riemann integration give Dedekind cuts naturally, but sequences artificially. Any Cauchy sequence with a specified modulus of convergence has a limit that is defined by a Dedekind cut, but it is more difficult to translate in the other direction. We suspect that the preference for Cauchy sequences is merely a symptom of the traditional prejudice against logic. This paper shows that Dedekind cuts can be defined without using set theory, and we hope to demonstrate in future work that they also provide a natural way in which to compute with real numbers [Bau08, K].

The Heine–Borel theorem (compactness of the closed interval) is one of the most important properties that real analysts use. However, as we shall see in Section 15, this is not just a result that we prove in passing, but a hotly debated issue in the foundations of constructive analysis. For example, Errett Bishop [Bis67] reformulated a large part of the subject without it, deftly avoiding the many pathologies that arise from so doing.

In particular, he defined a function to be "continuous" if it is *uniformly* so when restricted to any compact domain, "compact" itself being defined as closed and totally bounded. Hence any continuous $f : K \to \mathbb{R}^+$ is bounded (away from ∞, if you wish), but there is no similar result to bound it away from 0. Indeed, if K is Cantor space, such a result would entail (Brouwer's fan theorem and so) the Heine–Borel theorem. It follows that, if there is a category of partial functions on $\mathbb{R}$ that agrees with the uniformly continuous ones on compact domains and includes $\lambda x.1/x$, then the Heine–Borel theorem holds [Pal05, Sch05, Waa05, JR84, BB08b, BB08a]

The reason why Bishop's followers and others omit this property is that they are interested in *computation*, at least in principle. However, if one tries to develop analysis based on points, that is, in the way in which it has been done since Cantor, but using only those real numbers that are representable by programs, then the results are exceedingly unpleasant. In particular, the Heine–Borel theorem *fails.*

In fact, Dedekind completeness and the Heine–Borel property are both consequences of the view that *open sets and not points are primary.* That they hold at all in the traditional setting is the result of the heavy-handed methods of classical mathematics, which are far stronger than what is justified by computation. Brouwer's fan theorem is, arguably, a way of legitimising part of the classical approach in a constructive setting.

The best developed formulation of topology entirely in terms of open sets ("pointless topology", according to a now rather tired joke) is provided by locale theory. Although it does not consider computation, it does provide a way of developing general topology in foundational settings (at least, in toposes) other than the classical one. The most famous example of this is that it proves the Tychonov theorem (that the product of any family of compact objects is compact) without using the axiom of choice [Joh82, Theorem III 1.7]. Less well known, but more importantly for us, the *localic* interval $[0, 1]$ is always Dedekind-complete and compact. On the other hand, when we interpret the *traditional* (point-based) definitions in the internal language of a sheaf topos, the object of Cauchy reals is typically smaller than the Dedekind one, and the Heine–Borel theorem fails [FH79].

Formal Topology also works with open subspaces, but is based on Martin-Löf type theory; there too $[0, 1]$ is compact [CN96].

Abstract Stone Duality exploits the algebra of open sets as well, and so the "real line" R that we construct in this paper is Dedekind-complete and satisfies the Heine–Borel property. But ASD generalises Dedekind's topological conception of the real line: in it, the topology is an inherent and unalienable part of a space, which is not a set of points to which open subsets have been added as an afterthought.

In locale theory, the algebra and lattice theory are all too obvious, whilst they are represented in formal topology by generators and relations. Both of these theories expect a high degree of mathematical sophistication and perseverance from the student, and only in an exceptionally well written account do the public theorems about topology stand out from the private algebraic calculations. In [Joh82] the former are marked with an asterisk, although the official meaning of that symbol is a dependence on the axiom of choice.

The lattice of open subspaces in locale theory is a set (or an object of a topos), but in ASD it is another space, with its own topology. This is formulated in the style of a type theory that makes ASD *look* like topology with points. The λ-notation speaks out loud and clear about continuous maps in a way that frame homomorphisms in locale theory do not. When we do some basic analysis in [J], we shall see that the language of terms, functions and open predicates actually works more smoothly than does the traditional one using set theory.

In both traditional topology and locale theory there is an asymmetry between infinite unions and finite intersections that makes it difficult to see the duality between open and closed phenomena. Intuitionistic foundations also obscure this symmetry by stating many results that naturally pertain to closed sets in a form that uses double negations. When we treat the lattice of opens of one space as another space, by contrast, the purely infinitary (directed) joins slip into the background, and the open–closed duality stands out very clearly. Indeed, it is a fruitful technique to "turn the symbols upside down" ($\top/\bot$, $\wedge/\vee$, $=/\neq$, $\forall/\exists$), often giving a new theorem.

In this context, we shall see what the foundational roles of Dedekind completeness and the Heine–Borel theorem actually are. The former is the way in which the logical manipulation of topology has an *impact* on numerical computation. Again there is an analogy with the axioms for the *natural* numbers: for them the same role is played by *definition by description*, which Giuseppe Peano was also the first to formulate correctly [Pea97, §22], albeit in a different paper from the one on induction. In the ASD λ-calculus, these ideas are captured as rules that introduce numbers on the basis of logical premises.

The Heine–Borel property, meanwhile, is the result of ensuring that *all* algebras are included in the category. Our λ-calculus formulates this idea in an apparently point-based way by introducing higher-order terms which ensure that subspaces carry the subspace topology. We shall show that these terms are inter-definable with the "universal quantifiers" $\forall$ that define compactness.

The new concept of overtness is related to open subspaces in the way that compactness is to closed ones, and to logic in the shape of the existential quantifier, $\exists$. However, in contrast to compactness of the closed interval, no discipline would contest that the real line is overt. Indeed, the reason why you haven't heard of overt (sub)spaces before is that classical topology makes *all* spaces overt — by *force majeure*, without providing the computational evidence.

This idea has, in fact, been identified in locale theory, but only the experts in that subject have been able to exploit it [JT84, Joh84]. In the case of overtness, formal topology shows up the idea better than locale theory does. Its role in constructive analysis is played by *locatedness*, though that is a metrical property, so the correspondence is not exact [Spi07]. We shall show in ASD's account of analysis [J] that overtness explains the situations in which equations $fx = 0$ for $f : \mathbb{R} \to \mathbb{R}$ can or cannot be solved.

However, it is really in computation that the importance of this concept becomes clear. For example, it provides a generic way of solving equations, when this is possible.

Since ASD is formulated in a type-theoretical fashion, with absolutely no recourse to set theory, it is intrinsically a computable theory.

The familiar arithmetical operations $+$, $-$ are $\times$ are, of course, computable algebraic structure on R, as are division and the (strict) relations $<$, $>$ and $\neq$ when we introduce suitable types for their arguments and results. The topological properties of overtness and compactness are related to the logical quantifiers $\exists$ and $\forall$, which we shall come to see as additional computable structure.

Any term of the ASD calculus is in principle a program, although the details of how this might be executed have yet to be worked out [Bau08, K]. In particular, our proofs of compactness and overtness of closed intervals provide programs for computing quantifiers of the form $\forall x{:}[d, u]$ and $\exists x{:}[d, u]$ respectively. These are general and powerful higher-order functions from which many useful computations in real analysis can be derived.

This paper is rather long because we have to introduce the ASD calculus before we can use it for the construction. Although real *arithmetic* is familiar and not really related to the main issue of the Heine–Borel theorem, you would think it odd if we left it out, and of course we shall need it in order to prove completeness of the axioms, but it is a sizable task in itself. If you are impatient to see our construction of R, you may find it helpful to start with Section 6, and then bring in the introductory material as you need it.

We shall use a lot of *ideas* from interval analysis. However, instead of defining an interval $[d, u]$ as the set $\{x \in \mathbb{R} \mid d \le x \le u\}$ or as a pair $\langle d, u\rangle$ of real numbers, as is usually done, we see it as a weaker form of Dedekind cut, defined in terms of the rationals. Real numbers (genuine cuts) are special intervals. Section 2 explains the classical idea behind our construction of the Dedekind reals, presenting it from the point of view of interval analysis.

As this is the first paper in which Abstract Stone Duality has reached maturity, we give a survey of it in Sections 3–5 that will also be useful for reference in connection with other applications besides analysis. This provides a *guide* to the earlier papers, by no means making them redundant; for other independent introductions, see [J, O].

Sections 6–9 perform the main construction, developing cuts, the interval domain and the real line in ASD, and prove Dedekind completeness. Section 10 proves that the closed interval is compact and overt.

Sections 11–13 consider arithmetic in an entirely order-theoretic style, *i.e.* with Dedekind cuts rather than Cauchy sequences. We formulate an important "roundedness" property of the interval operations that is crucial to their correctness, extend the arithmetic operations to Kaucher's back-to-front intervals, and identify the precise role of the Archimedean principle.

After we have shown how to construct *an* object that satisfies Definition 1.1, Section 14 shows that it is unique (up to unique isomorphism), *i.e.* that the axioms above are complete. This means, on the one hand, that they are sufficient to develop analysis [J], and on the other that we may focus on them in order to do computation. We also show that Dedekind completeness is equivalent to abstract sobriety *via* some simple λ-conversions.

Finally, Section 15 compares ours with other schools of thought. In particular, we contrast compactness of the closed interval here with its pathological properties in Recursive Analysis, and comment on the status of ASD from the point of view of a constructivist in the tradition of Errett Bishop.

2 Cuts and intervals

We begin by recalling Richard Dedekind's construction, and relating it to some ideas in interval analysis. These provide the classical background to our abstract construction of R in ASD, which will start with our formulation of Dedekind cuts in Section 6. In this section we shall *use* the Heine–Borel theorem to prove an abstract property of $\mathbb{R}$, from which we deduce the fundamental theorem of interval analysis. The same abstract property, taken as an *axiom*, will be the basis of our construction of R and proof of the Heine–Borel theorem in ASD.

Remark 2.1 Dedekind [Ded72] represented each real number $a \in \mathbb{R}$ as a pair of sets of rationals, $\{d \mid d \le a\}$ and $\{u \mid a < u\}$. This asymmetry is messy, even in a classical treatment, so it is better to use the strict inequality in both cases, omitting a itself if it's rational. So we write

$$D_a \;\equiv\; \{d \in \mathbb{Q} \mid d < a\} \quad \text{and} \quad U_a \;\equiv\; \{u \in \mathbb{Q} \mid a < u\}.$$

These are disjoint inhabited open subsets that "almost touch" in the sense that

$$d < u \implies d \in D \;\vee\; u \in U.$$

This property is usually known as *locatedness*.

The idea that a set of *rationals* is open in the usual topology on $\mathbb{R}$ and extends towards $\pm\infty$ can be expressed order-theoretically, using a condition that we shall call *roundedness*:

$$d \in D \iff \exists e.\, d < e \wedge e \in D \qquad \text{and} \qquad u \in U \iff \exists t.\, t < u \wedge t \in U.$$

A real number represented as a cut is therefore a pair of subsets, *i.e.* an element of $\mathcal{P}(\mathbb{Q}) \times \mathcal{P}(\mathbb{Q})$, that is rounded, bounded (*i.e.* inhabited), disjoint and located.

Exercise 2.2 If $\Delta \subset \mathbb{R}$ is an open set of reals that extends downwards, show that $D \equiv \Delta \cap \mathbb{Q}$ is a rounded set of rationals with the same property, and that this defines a bijection between such subsets, where $\Delta \equiv \{r \in \mathbb{R} \mid \exists q \in D.\, r < q\}$. In future we shall write D for either representation, and pass between them without comment; in particular, in "$\mathbb{R} \setminus (D \cup U)$" below, real as well as rational numbers are removed.

Definition 2.3 It is very useful to generalise the notion of Dedekind cut by dropping the locatedness condition. Instead of almost touching, and so representing a *single* real number $a \in \mathbb{R}$, such a ***pseudo-cut*** corresponds classically to a *closed interval* $[d, u] \equiv \mathbb{R} \setminus (D \cup U)$. We sometimes weaken the other conditions too, allowing $D \equiv \emptyset$ and $d \equiv -\infty$, or $U \equiv \emptyset$ and $u \equiv +\infty$.

The extension of the arithmetic operations to such intervals was defined by Ramon Moore [Moo66]:

$$\begin{array}{lcll}
[d,u] \oplus [e,t] & \equiv & [d+e, u+t] & \\
\ominus [d,u] & \equiv & [-u, -d] & \\
[d,u] \otimes [e,t] & \equiv & [\mathsf{min}(de, dt, ue, ut),\ \mathsf{max}(de, dt, ue, ut)] & \\
[d,u]^{-1} & \equiv & [u^{-1}, d^{-1}] & \text{if } 0 \notin [d,u], \text{ so } 0 \in D \cup U \\
 & \equiv & [-\infty, +\infty] & \text{if } 0 \in [d,u].
\end{array}$$

The formula for multiplication is complicated by the need to consider all possible combinations of signs.

Remark 2.4 Moore's ***interval analysis*** has been used to develop a variety of numerical algorithms; see [Kea96] for a brief survey with an extensive bibliography. Amongst these, we focus on what it achieves for the problem of *optimisation*. By this, we understand finding the *maximum value* of a continuous function defined on a non-empty compact domain, but not necessarily any *location* where the function attains that value. Plainly, *any* value of the function provides a *lower* bound for the maximum, but finding *upper* bounds is problematic using standard numerical methods, especially when the function has "spikes".

For the sake of illustration, consider an arithmetical function $f : [0,1]^n \to \mathbb{R}$. If this is just addition or multiplication, Moore's interval operations provide the minimum and maximum values of the function on the domain directly. For a more complicated arithmetical function, we interpret the operations according to Moore's formulae, and may (if we're lucky) still obtain the minimum and maximum.

In general, however, the result of Moore's interpretation will be an interval that *contains* the required image. In other words, it provides an *upper* bound of the maximum — exactly what standard numerical methods find difficult. Unfortunately, this may be a *vast* over-estimate, especially when the definition of the function is complicated and involves subtraction of similar large numbers, so this does not really help very much with spiky functions.

In fact, Moore and his followers have provided various techniques for reducing the sizes of the resulting intervals. One of these simply massages the arithmetic expression to reduce multiple occurrences of variables and sub-expressions, since computing $x - x$ introduces a large error that can easily be avoided. However, these techniques are not the purpose of the present discussion.

Remark 2.5 We regard Moore's definitions as merely *one* way of extending certain continuous functions to take intervals instead of real numbers as arguments. In fact, that's exactly the point:
(a) *ideally*, we extend f to the operation F_0 that takes intervals to intervals by

$$F_0[d,u] \equiv \{fx \mid x \in [d,u]\},$$

(b) but *in practice*, Moore's operations extend a general arithmetic expression f to *some* continuous operation F on intervals such that

$$F[a,a] \;=\; [fa, fa] \quad \text{and} \quad F_0[d,u] \subset F[d,u].$$

Whilst Moore's interpretation may give a huge over-estimate of the image if we don't rewrite the expression, it has an important technical property, namely that *it preserves (syntactic) composition of arithmetic expressions*, so it is easily performed by a compiler.

The first question that this raises is whether the ideal operation F_0, which (a) defines set-theoretically, is *continuous*. Then we can ask whether there is a way of *computing* it, using (b).

Continuity is crucial to our whole development: not only that of the single operation F_0, but also for the process that extends f to F or F_0. However, as $\mathcal{P}(\mathbb{Q})$ is a complete lattice, rather than a space that has a familiar Hausdorff topology motivated by geometry, we first have to describe what its *topology* is. (In this paper, we do not even attempt to put a topology on $\mathbb{R}^{\mathbb{R}}$.)

Definition 2.6 Let L be a complete lattice such as $\mathcal{P}(\mathbb{Q})$. Then a family $\mathcal{U} \subset L$ is called ***Scott open*** [Sco72a, GHK+80] if it is
(a) an upper set, so if $U \in \mathcal{U}$ and $U \leqslant V$ then $V \in \mathcal{U}$; and
(b) "inaccessible by directed unions", *i.e.* if $(\bigvee_{i\in I} U_i) \in \mathcal{U}$ then already $(\bigvee_{i\in F} U_i) \in \mathcal{U}$ for some finite subset $F \subset I$.
This ***Scott topology*** is never Hausdorff (except, that is, on the trivial lattice).

A function $f : L_1 \to L_2$ between complete lattices is continuous with respect to the Scott topology iff it preserves directed unions. In particular, it is monotone, *i.e.* if $x \leqslant y$ in L_1 then $fx \leqslant fy$ in L_2. See [J] for a brief account of this topology that is enough for its use here.

Exercise 2.7 Let L be the topology (lattice of open subsets) of any topological space X, and $K \subset X$ any subset. Show that the family $\mathcal{U} \equiv \{U \mid K \subset U\}$ of open neighbourhoods of K is Scott open iff K is ***compact*** in the usual "finite open sub-cover" sense. If this is new to you, you would be well advised to stop *at this point* to see for yourself why this is the case, since this idea is fundamental to our whole programme. □

Example 2.8 Let $L \equiv \underline{\mathbb{R}} \equiv \mathbb{R} + \{-\infty, +\infty\}$, considered as a complete lattice in the arithmetical order, so $+\infty$ is the top element. Equipped with the Scott topology, this space arises in real analysis because ***lower semicontinuous*** functions $X \to \mathbb{R}$ are just functions $X \to \underline{\mathbb{R}}$ that are continuous in the sense of general topology. We shall see, however, that it is a mistake to think of $\underline{\mathbb{R}}$ as *derived* from $\mathbb{R}$ in this clumsy way: it is really simpler and more fundamental than $\mathbb{R}$ itself. Its elements, which we call ***ascending reals***, are given by Exercise 2.2, *i.e.* in a similar way to the Dedekind reals, except that only the lower cut $D \subset \mathbb{Q}$ is used. The ***descending reals***, $\overline{\mathbb{R}}$, are defined in the same way, but with the opposite order, and are related to *upper* semicontinuous functions. The words "ascending" and "descending" relate the usual temporal language about the real line to the vertical intuition of lattice theory. □

Remark 2.9 It is useful theoretically to consider more general intervals than Moore did, *i.e.* not necessarily having rational endpoints. For example, a computation that is allowed to continue forever may generate a converging pair of sequences of rationals,

$$d_0 < d_1 < d_2 < \cdots < u_2 < u_1 < u_0,$$

whose limits $\mathsf{sup}\, d_n$ and $\mathsf{inf}\, u_n$ we would like to see as the 'ultimate result' of the computation.

If these limits exist in $\mathbb{R}$, say $d \equiv \mathsf{sup}\, d_n$ and $u \equiv \mathsf{inf}\, u_n$, then $[d, u]$ is the (directed) intersection $\bigcap [d_n, u_n]$. However, d and u are best seen as *ascending* and *descending reals*,

$$D \ \equiv\ \{d \mid \exists n.\, d < d_n\} \in \underline{\mathbb{R}} \qquad \text{and} \qquad U \ \equiv\ \{u \mid \exists n.\, u > u_n\} \in \overline{\mathbb{R}},$$

since, constructively, they need not exist as members of the usual "Euclidean" structure (Remark 15.1). Nevertheless, the intersection of the closed intervals does exist: it is $\mathbb{R} \setminus (D \cup U)$, and is always a *closed* subspace.

Notice that the *open* subsets D and U represent *positive* information, namely the lower and upper bounds that we have verified *so far* for the real number that we're trying to calculate. The *closed* interval $[d, u]$, on the other hand, consists of the candidate real numbers that we have *not yet excluded*, which is *negative* information.

Definition 2.10 We can consider these generalised intervals as members of the ***interval domain***, $\mathbb{IR}$ [Sco72b, ES98]. The order relation is traditionally defined as *reverse* inclusion of closed intervals, and so directed joins are given by *intersection*.

According to our preferred view, an interval is defined as a generalised Dedekind cut (D, U), and the order relation is the *forward* inclusion of these sets. As we acquire more information about the number, in the form of tighter bounds, the correspondingly narrowing interval goes *up* in the order on the domain.

We shall now start using some of the notation of the calculus that we intend to introduce.

Notation 2.11 Classically, any subset $D \subset \mathbb{Q}$ may be represented as a function $\mathbb{Q} \to \{\top, \bot\}$ that says whether or not each $d \in \mathbb{Q}$ belongs to D. Since we believe that topology is fundamental, we give this target set the Scott topology. It then becomes the ***Sierpiński space***, in which just one singleton $\{\top\}$ is open, making the other $\{\bot\}$ closed. It therefore looks like $\left(\begin{smallmatrix}\odot\\ \bullet\end{smallmatrix}\right)$, and we shall call it Σ.

Since $\mathbb{Q}$ is discrete, it doesn't change much to require our functions $\mathbb{Q} \to \Sigma$ to be continuous. However, for a general topological space X (such as $\mathbb{R}$), continuous functions $\phi : X \to \Sigma$ correspond bijectively to *open* subspaces $U \subset X$, by $U \equiv \phi^{-1}(\top)$. We shall say that ϕ ***classifies*** U.

Notation 2.12 Having passed from the (open) subset $D \subset \mathbb{Q}$ to a (continuous) function $\delta : \mathbb{Q} \to \Sigma$, we also replace set-theoretic notation with the λ-calculus. A function $x \mapsto \phi(x)$ is written as a λ-*abstraction* $\lambda x.\, \phi(x)$; we use nested λ-abstractions for several variables, but write $\lambda xy.\, \sigma$ instead of $\lambda x.\, \lambda y.\, \sigma$. Application $\phi(a)$ is written without parentheses, ϕa. Expressions built from λ-abstractions and applications are called λ-*terms*.

So in ASD we represent D and U by λ-terms $\delta, \upsilon : \Sigma^{\mathbb{Q}}$, adopting the convention of using Greek letters for terms of this type. Many authors in computer science would write $\mathbb{Q} \to \Sigma$ instead of $\Sigma^{\mathbb{Q}}$, which is the notation in pure mathematics. Membership $d \in D$ is written δd, and set-formation becomes λ-abstraction. In particular, D_a and U_a become

$$\delta_a \ \equiv\ \lambda d.\, (d < a) \quad \text{and} \quad \upsilon_a \ \equiv\ \lambda u.\, (a < u),$$

so $\delta_a d$ means $d < a$. Then any real number can be represented as a pair (δ, υ) of λ-terms, and the real line as a whole is a subset $\mathbb{R} \subset \Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}$.

Remark 2.13 We want to use this representation to compute with real numbers, and in the first instance to do arithmetic with them. Dedekind indicated how this can be done, defining operations on cuts. But there is a difference between his objective of providing a rigorous foundation for differential and integral calculus, and ours of getting a machine to compute with real numbers. He only had to define and justify the operations on *legitimate* cuts (those that are rounded, bounded, disjoint and located), whereas our machine will do *something* with *any* pair of λ-terms that we give it, even if that's only to print an error message.

It's reasonable to ask that any program F that is intended to compute a function $f:\mathbb{R}\to\mathbb{R}$ using Dedekind cuts will actually take *any* pair of λ-terms and return another such pair. We say that the program is *correct* for this task if, when it is given the pair (δ_a,υ_a) that represents a number a, it returns the pair $(\delta_{fa},\upsilon_{fa})$ that represents the value $f(a)$ of the function at that input. In other words, the square on the left commutes:

$$\begin{tikzcd} \mathbb{R} \arrow[r, tail, "i"] \arrow[d, "f"'] & \Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}} \arrow[d, dotted, "F"] \\ \mathbb{R} \arrow[r, tail, "i"] & \Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}} \end{tikzcd} \qquad \begin{tikzcd} \mathbb{R} \arrow[r, tail, "i"] \arrow[d, "\phi"'] & \Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}} \arrow[dl, dotted, "\Phi"] \\ \Sigma & \end{tikzcd}$$

Remark 2.14 We shall not attack the problem of extending functions directly. It is simpler and more in keeping with the ideas of computation, Dedekind cuts and topology to consider first how an *open subset* of $\mathbb{R}$ may be extended to one of $\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}$.

Recall from Notation 2.11 that such an open subset is classified by a continuous function taking values in Σ. We expect the function ϕ from $\mathbb{R}$ (classifying an open subset of it) to be the restriction of a function Φ from $\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}$ (classifying an open subset of that), making the triangle on the right above commute. In other words, $\mathbb{R}$ should carry the ***subspace topology*** inherited from the ambient space, $\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}$, which itself carries the Scott topology.

This extension of open subspaces is actually the fundamental task, since we can use it to extend functions by defining

$$F(\delta,\upsilon)\ \equiv\ \big(\lambda d.\ \Phi_d(\delta,\upsilon),\ \lambda u.\ \Psi_u(\delta,\upsilon)\big),$$

where Φ_d and Ψ_u are the extensions of $\phi_d\equiv\lambda x.\,(d<fx)$ and $\psi_u\equiv\lambda x.\,(fx<u)$ respectively.

The extension process must therefore respect parameters. We implement this idea by introducing a new operation I that extends ϕ to Φ in a *continuous, uniform* way, instead of a merely existential property of the extension of functions or open subspaces *one at a time.* This depends on an understanding of Scott-continuous functions of higher type.

Proposition 2.15 Classically, there is a map $I:\Sigma^{\mathbb{R}}\rightarrowtail\Sigma^{\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}}$, defined by

$$(V\subset\mathbb{R})\text{ open }\ \mapsto\{(D,U)\mid \exists d\in D.\,\exists u\in U.\,(d<u)\wedge[d,u]\subset V\}$$

in traditional notation, or

$$\phi:\Sigma^{\mathbb{R}}\mapsto\lambda\delta\upsilon.\ \exists du.\ \delta d\wedge\upsilon u\wedge(d<u)\wedge\forall x{:}[d,u].\ \phi x$$

in our λ-calculus, that is *Scott-continuous.* It makes $\Sigma^{\mathbb{R}}$ a retract of $\Sigma^{\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}}$, as it satisfies the equation

$$\Sigma^i\cdot I\ =\ \mathsf{id}_{\Sigma^{\mathbb{R}}}\qquad\text{or}\qquad x:\mathbb{R},\ \phi:\Sigma^{\mathbb{R}}\ \vdash\ I\phi(ix)\ \Leftrightarrow\ \phi x,$$

where $ix\equiv(\delta_x,\upsilon_x):\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}$. This expresses *local compactness* of $\mathbb{R}$.

Proof The Heine–Borel theorem, *i.e.* the "finite open sub-cover" definition of compactness for the closed interval $[d,u]$, says exactly that the expression "$[d,u]\subset V$" is a Scott-continuous predicate in the variable $V : \Sigma^{\mathbb{R}}$ (Exercise 2.7). Thus the whole expression for I is a Scott-continuous function of V. This satisfies the equation because, if ϕ classifies $V \subset X$,

$$\phi a \;\equiv\; (a \in V) \;\mapsto\; I\phi(ia) \;\equiv\; \big(\exists du.\, a \in (d,u) \subset [d,u] \subset V\big) \iff (a \in V). \qquad \square$$

The argument so far has relied on the prior existence of $\mathbb{R}$. The next step is to eliminate real numbers in favour of rational ones.

Lemma 2.16 Since $\Phi : \Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}} \to \Sigma$ is a Scott-continuous function,

$$\Phi(\delta_f, \upsilon_s) \iff \exists et.\, (e < f) \land (s < t) \land \Phi(\delta_e, \upsilon_t) \qquad \text{(a)}$$

$$\Phi(\delta_x, \upsilon_x) \iff \exists mk.\, (p_{k-1} < x < p_{k+1}) \land \Phi(\delta_{p_{k-1}}, \upsilon_{p_{k+1}}) \qquad \text{(b)}$$

where $p_k \equiv d + 2^{-m}k(u-d)$, for any chosen $d < u$. Although p_k also depends on m, d and u, we omit them to simplify the notation: they should be understood to come from the expression in which p_k is embedded.

Proof In traditional notation, the unions

$$D_f \;=\; \bigcup_{e<f}^{\uparrow} D_e \quad \text{and} \quad U_s \;=\; \bigcup_{s<t}^{\uparrow} U_t$$

are directed, *cf.* Remark 2.9, so Φ preserves them. The equations above say the same thing in our λ-calculus (Notation 2.12). Notice that $(\Leftarrow)$ is monotonicity. Having enclosed the point x in some open interval (e,t), we may find m and k so that $e < p_{k-1} < x < p_{k+1} < t$. $\square$

Proposition 2.17 The idempotent $\mathcal{E} \equiv I \cdot \Sigma^i$ on $\Sigma^{\Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}}$ is given by

$$\mathcal{E}\Phi(\delta,\upsilon) \;\equiv\; \exists n \geq 1.\, \exists q_0 < \cdots < q_n.\;\; \delta q_0 \land \upsilon q_n \;\land\; \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k},\; \upsilon_{q_{k+1}}).$$

Proof Substituting $ix \equiv (\delta_x, \upsilon_x)$ into the definition of I in Proposition 2.15,

$$(I \cdot \Sigma^i)\Phi(\delta,\upsilon) \;\equiv\; I\big(\lambda x.\, \Phi(ix)\big)(\delta,\upsilon) \;\equiv\; \exists du.\, \delta d \land \upsilon u \land (d < u) \land \forall x{:}[d,u].\, \Phi(\delta_x, \upsilon_x).$$

$[\mathcal{E} \leqslant I \cdot \Sigma^i]$: Although the formula for $\mathcal{E}$ essentially involves *abutting closed* intervals,

$$[d,u] \;=\; [q_0,q_1] \cup [q_1,q_2] \cup \cdots \cup [q_{n-1},q_n],$$

part (a,$\Rightarrow$) of Lemma 2.16 expands each of them slightly, from $[q_k, q_{k+1}]$ to $[e,t]$. Then each $x \in [d,u]$ lies inside one of the *overlapping open* intervals (e,t):

$$\begin{aligned}
\mathcal{E}\Phi(\delta,\upsilon) \;&\Rightarrow\; \exists q_0 \ldots q_n.\;\; \delta q_0 \land \upsilon q_n \;\land\; \bigwedge_{k=0}^{n-1} \exists et.\, (e < q_k < q_{k+1} < t) \;\land\; \Phi(\delta_e, \upsilon_t) \\
&\Rightarrow\; \exists du.\;\; d < u \;\land\; \delta d \land \upsilon u \;\land\; \forall x{:}[d,u].\, \exists et.\, (e < x < t) \;\land\; \Phi(\delta_e, \upsilon_t) \\
&\Rightarrow\; \exists du.\;\; d < u \;\land\; \delta d \land \upsilon u \;\land\; \forall x{:}[d,u].\, \Phi(\delta_x, \upsilon_x),
\end{aligned}$$

where the last step uses Lemma 2.16(a,$\Leftarrow$) with $f \equiv s \equiv x$.

$[\mathcal{E} \geqslant I \cdot \Sigma^i]$: For the converse, Lemma 2.16(b,$\Rightarrow$) also encloses each point $x \in [d,u]$ in an open interval $(p_{k-1}, p_{k+1}) \ni x$ with dyadic endpoints, such that $\Phi(p_{k-1}, p_{k+1})$ holds. Although some points x may *â priori* need narrower intervals than others, with larger values of m, the Heine–Borel theorem says that finitely many of them suffice, so there is a *single* number m that serves for *all* $x \in [d,u]$. Another way of saying this is that the universal quantifier $\forall x : [d,u]$ is Scott-continuous, preserving the directed join $\exists m$.

$$\begin{aligned}
&\forall x{:}[d,u].\ \Phi(\delta_x, \upsilon_x) \\
&\qquad \Rightarrow\quad \forall x{:}[d,u].\ \exists m{:}\mathbb{N}.\ \exists k.\ (p_{k-1} < x < p_{k+1})\ \wedge\ \Phi(\delta_{p_{k-1}},\ \upsilon_{p_{k+1}}) \\
&\qquad \Rightarrow\quad \exists m.\ \forall x{:}[d,u].\ \exists k.\ (p_{k-1} < x < p_{k+1})\ \wedge\ \Phi(\delta_{p_{k-1}},\ \upsilon_{p_{k+1}}) \\
&\qquad \Rightarrow\quad \exists m.\ \bigwedge_{k=0}^{2^m} \Phi(\delta_{p_{k-1}},\ \upsilon_{p_{k+1}}) \quad\Rightarrow\quad \exists m.\ \bigwedge_{k=0}^{2^m-1} \Phi(\delta_{p_k},\ \upsilon_{p_{k+1}}),
\end{aligned}$$

where we trim the intervals using Lemma 2.16(a,$\Rightarrow$) and omit the last one. Then $n \equiv 2^m$, $q_0 \equiv p_0 \equiv d$, $q_1 \equiv p_1, \ldots, q_n \equiv p_{2^m} \equiv u$ provide a sequence to justify $\mathcal{E}\Phi(\delta, \upsilon)$. □

Remark 2.18 Notice that this formula for $\mathcal{E}$ involves only rational numbers (and predicates on them) *but not reals*. We can *write* this formula with very little assumption about the underlying logic. The role of the classical Heine–Borel theorem was to define a different map, I, to prove that it is Scott-continuous and that it satisfies the equations

$$\Sigma^i \cdot I \;=\; \mathsf{id}_{\Sigma^{\mathbb{R}}} \quad \text{and} \quad I \cdot \Sigma^i \;=\; \mathcal{E}.$$

It is the formula $\mathcal{E}$ — not the classically defined map I or the accompanying proof — that we shall use to construct the real line in ASD. In Section 15 we shall discuss some systems of analysis in which the Heine–Borel theorem *fails*, and the operation I does not exist either.

Besides the construction of R in ASD, the formula is important for many other reasons. In particular, it shows how Moore's interval arithmetic solves the optimisation problem, and justifies the interval-subdivision algorithms that have been developed using interval analysis. It would be appropriate to call it the ***fundamental theorem of interval analysis***. Moore himself proved it [Moo66, Chapter 4] using convergence with respect to the metric

$$\partial\big([d,u],[e,t]\big) \;\equiv\; \mathsf{max}\,\big(|d-e|, |u-t|\big).$$

Corollary 2.19 Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function and F be *any* continuous extension of f to intervals (Remark 2.13). For example, if f is an arithmetic expression then F may be its interpretation using Moore's operations. Suppose that e and t are strict bounds for the image of $[d,u]$ under f, so

$$\forall x{:}[d,u].\ e < fx < t, \quad \text{or, equivalently,} \quad F_0[d,u] \subset (e,t),$$

where F_0 gives the set-theoretic image. Then there is some finite sub-division (which could be chosen to be dyadic), $d \equiv q_0 < q_1 < \cdots < q_{n-1} < q_n \equiv u$, such that

$$(e,t) \;\supset\; \bigcup_{m=0}^{n-1} F[q_k, q_{k+1}].$$

Moreover, we obtain F_0 from F in Remark 2.5 using $\mathcal{E}$ and Remark 2.14. □

This sub-division could be found by a recursive program that either applies the function Moore-wise to the whole interval, or splits it and calls itself on each half.

Remark 2.20 Finally, some curious things happen in both interval analysis and ASD when we play around with the formulae. Of course, if $a < b$ then the subsets

$$D \equiv D_b \equiv \{d \mid d < b\} \quad \text{and} \quad U \equiv U_a \equiv \{u \mid a < u\}$$

overlap. But we may still apply Proposition 2.15 to them, obtaining the *existential* quantifier,

$$(D_b, U_a) \in IV \iff \exists x{:}[a,b].\ x \in V.$$

We shall see in Section 10 that both quantifiers are derived from $\mathcal{E}$ like this.

Given that $\forall$ is related to *over*estimation of the range of a function in interval analysis, could this back-to-front version for $\exists$ yield an *under*estimate? It is obviously very dangerous in this case to rely on any intuition of an interval as a single *closed* set, as is common in the literature on interval analysis. On the other hand, there is no problem in dealing with a back-to-front interval in our formulation of it as a *pair* (D, U) of (now overlapping) open sets.

In order to exploit this idea, we would have to begin with a careful study of the arithmetic, which was introduced by Edgar Kaucher [Kau80]. Addition and subtraction are as in Remark 2.3, but multiplication is more complicated (Lemma 12.1). Various authors have used Kaucher's arithmetic to investigate questions of linear algebra, logic and underestimation, *e.g.* [KNZ96, Lak95].

3 Topology as lambda-calculus

Now we shall introduce Abstract Stone Duality. It is a direct axiomatisation of general topology whose aim is to integrate it with computability theory and denotational semantics, without sacrificing important properties such as the Heine–Borel theorem. The basic building blocks — spaces and maps — are taken as fundamental, rather than being manufactured from sets by imposing extra structure.

The ASD calculus formalises the notation that we have already used for the set-theoretic constructions in the previous section (Notation 2.11ff):

(a) the Sierpiński space Σ and its lattice structure;

(b) λ-terms of type Σ^X to encode open subspaces of X;

(c) $\exists$ for combinatorial structure;

(d) $\forall$ to say when the direct image of a closed bounded interval is contained in an open interval;

(e) real numbers represented as Dedekind cuts, considered as pairs (δ, υ) of λ-terms; and

(f) the Σ-splitting I that is related to compactness of the interval $[d, u]$.

As a type theory, ASD has *types*, *terms* and *equations* between terms. The types are also called *spaces*, and the terms *maps* or *functions*. The three basic spaces $\mathbf{1}$, $\mathbb{N}$ and Σ are axiomatised by their universal properties. From them, we may construct products, $X \times Y$, and exponentials of the form Σ^X, but *not* arbitrary Y^X. The theory also provides certain "Σ-split" subspaces, and it is our main objective to construct R in this way, using the idea of Proposition 2.15.

As Notation 2.11 suggests, terms of type Σ^X behave very much like predicates on X, and, more basically, those of type Σ like propositions. We can form conjunctions and disjunctions of such terms, but not implications or negations — these become equations *between* terms. In some cases there are also operators $\exists_X : \Sigma^X \to \Sigma$ and $\forall_X : \Sigma^X \to \Sigma$ that satisfy the same formal properties as the quantifiers (*cf.* Remark 4.16). A space X is called *compact* when it has a $\forall_X$ operator, and *overt* when it has $\exists_X$.

Statements in the theory are expressed as equations between terms, but we use $\leqslant$ or $\Rightarrow$ as syntactic sugar in the lattices Σ^X and Σ. Since not every formula of (for example) the predicate calculus is of this form, the theory imposes limitations on what can be said within it.

You may think that we are making a rod for our own backs by restricting ourselves to such a weak calculus. This criticism puts us in good company with constructive mathematicians, who often face the same lack of understanding. Indeed, one of the reasons why the papers on ASD are so long is that proofs in weak calculi — where they exist at all — necessarily have far more steps than those in stronger systems: it's like digging a trench with a teaspoon.

But the ASD calculus *is* the calculus of spaces and maps. If some feature is missing from it, this is not because of our asceticism, but because spaces and maps do not possess it. The justification of this claim is that, starting from the axioms of ASD, we may reconstruct the categories of computably based locally compact locales [G] and of general locales over an elementary topos [H].

In such a weak calculus, it will not surprise you to hear that, as a rule, it is very difficult to know how to say anything at all. But experience has shown something very remarkable, namely that, once we have *some* way of expressing an idea, it usually turns out to be the *right* way. In contrast, stronger calculi, *i.e.* those that take advantage of the logic of sets, type theory or a topos, offer numerous candidates for formulating an idea, but these then often lead to distracting counterexamples.

It will be useful to state the formal relationship between ASD and traditional topology.

Definition 3.1 In the ***classical interpretation*** of ASD, each type X denotes a locally compact topological space. This space need not be Hausdorff, so by "local compactness" we mean that each $x \in U \subset X$ has some $x \in V \subset K \subset U \subset X$, where U and V are open and K is compact [GHK+80]. Similarly, each term $x : X \vdash fx : Y$ denotes a continuous function $f : X \to Y$. All maps between spaces are continuous, simply because the calculus never introduces discontinuous functions.

The exponential Σ^X is the topology (lattice of open subspaces) of X, regarded not as a set but as another topological space, equipped with the Scott topology (Definition 2.6); since X is locally compact, Σ^X is a continuous lattice. Maps $\Sigma^X \to \Sigma^Y$ are Scott-continuous functions. The connectives $\top$, $\bot$, $\wedge$ and $\vee$ are interpreted in the obvious way in the lattice, and $\exists_N$ by N-indexed unions or joins. When K is a compact space, $\forall_K : \Sigma^K \to \Sigma$ denotes the Scott-continuous function in Exercise 2.7.

Remark 3.2 We usually think of a map $\phi : X \to \Sigma$ as representing an *open* subspace of X, namely the inverse image of the open point $\top \in \Sigma$. But, by considering the inverse image of the closed point $\bot \in \Sigma$ instead, the same term also corresponds to a *closed* subspace.

Hence there is a bijection between open and closed subspaces, but it arises from their common classifiers, *not* by complementation. Indeed, *there is no such thing as complementation, as there is no underlying theory of sets* (or negation).

This extensional correspondence between terms and both open and closed subspaces gives rise to Axiom 4.7 and Lemma 4.9, which are known as the ***Phoa principle***. This makes the difference between a λ-*notation* for topology that must still refer to set theory to prove its theorems, and a *calculus* that can prove theorems for itself, indeed in a computational setting. It was the subject of the first published paper on ASD [C].

Constructive or intuitionistic logicians may, however, feel uncomfortable using this axiom, because it *looks* like classical logic. Nevertheless, the "classical" interpretation above is also valid for *intuitionistic* locally compact locales (**LKLoc**) over any elementary topos.

The reason for this is that ASD is a higher-order logic of *closed* subspaces, just as much as it is of *open* ones. When it is functioning in its "closed" role, we are essentially using the symbols the wrong way round: for example, $\vee$ denotes intersection.

The discomfort will be particularly acute in this paper, as it deals with a Hausdorff space, where the diagonal $R \subset R \times R$ is closed (Definition 4.12). This means that we can discuss equality of two real numbers, $a, b : R$, using a predicate of type Σ. But since the logic of closed predicates is upside down, the (open) predicate is $a \neq_R b$, and for equality we have to write the equation $(a \neq b) \Leftrightarrow \bot$ between terms of type Σ. For the integers, on the other hand, we can state equality as $(n = m) \Leftrightarrow \top$.

That equality of real numbers is "doubly negated" like this was recognised by Brouwer, *cf.* [Hey56], and arguably even by Archimedes. However, we stress that this is not what we are doing in ASD: $\neq$ is primitive, not derived from equality, and there is no negation of predicates. The equational statements $(-) \Leftrightarrow \top$ and $(-) \Leftrightarrow \bot$ merely say that the parameters of the expression $(-)$ belong respectively to certain open or closed subspaces. These are of the same status in the logic.

Remark 3.3 As we shall see, it is the part of the calculus that concerns *(sub)types* that is crucial to the compactness of the closed interval.

There are many approaches to topology that are based on sets of points, although the word "set" may be replaced by (Martin–Löf) *type* or *object* (of a topos), and "functions" may be accompanied by *programs*. In all of these systems, a new object X is constructed by specifying a predicate on a suitable ambient object Y, and then carving out the desired subobject $i : X \rightarrowtail Y$ using the subset-forming tools provided by the foundations of the system. For example, the real line may be constructed as the *set* (type, object) of those pairs (D, U) or (δ, υ) that satisfy the properties required of a Dedekind cut.

What is the topology on such a subset? We fall on the mercy of the underlying logic of sets of points to determine this, and so to prove important theorems of mathematics such as compactness of the real closed interval. As we shall see in Section 15, in many interesting logical systems, especially those that are designed to capture recursion, the logic may not oblige us with a proof, but may yield a counterexample instead.

We consider that this *topological* problem is not about the underlying *logic* but about the existence of certain *continuous functions* of higher type. We saw in Proposition 2.15 that, in the classical situation, we may use the Heine–Borel theorem to define a Scott-continuous map I that extends any open subset of $X \equiv \mathbb{R}$ to one of $Y \equiv \Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}$. The other foundational systems in which Heine–Borel is provable can also define this map, whilst others do neither.

The key feature of Abstract Stone Duality is that it provides the map $I : \Sigma^X \rightarrowtail \Sigma^Y$ (called a Σ-splitting) *axiomatically* whenever we define a subspace $i : X \rightarrowtail Y$, thereby taking account of the intended topology. It is important to understand that the types that are generated by the language that we describe are therefore *abstract* ones, not sets of points.

The essence of our construction in the rest of the paper is then that the quantifier $\forall_{[0,1]}$ that makes $[0, 1]$ compact may conversely be derived from this map I.

Beware, however, that there are some subspaces (equalisers, Definition 5.3) that have no Σ-splitting in ASD or any other system of topology. For example, $i : \mathbb{N}^{\mathbb{N}} \rightarrowtail \Sigma^{\mathbb{N} \times \mathbb{N}}$ can be expressed as an equaliser of topological spaces or locales, but it does not have a *Scott-continuous* Σ-splitting I. The direct image map, called i_*, does satisfy the relevant equation, $i^* \cdot i_* = \mathsf{id}$, where i^* is localic notation for our Σ^i, but i_* is only *monotone*, not Scott-continuous. Consequently, $\mathbb{N}^{\mathbb{N}}$ is not definable in ASD as the calculus is described here. On the other hand, a particular inclusion may have many different Σ-splittings, as we shall find in Section 7.

We formalise the subspace calculus in Section 5. However, as we shall see in Sections 7 and 8, it is not as easy to use as set-theoretic comprehension — the rule has been to introduce only one subspace per paper. So far, it only gives a robust account of locally compact spaces, but there is work in progress to extend it.

Remark 3.4 Since ASD is presented as a λ-calculus, it is reasonable to ask whether
(a) it is powerful enough for us to write programs in it;
(b) it has some computational interpretation, at least in principle; and
(c) we can actually print out digits of real numbers using it.
Each of these questions deserves a paper in itself. You will find several remarks here that are motivated by practical computation, and we are actively thinking about a prototype implementation [Bau08], but we do not say very much about it in this paper.

The language was indeed partly inspired by ideas from theoretical computer science. Computation is a natural part of the calculus, and there is no need to bolt a clumsy, old-fashioned theory of recursion on to the front of it. Domain theory can be developed within ASD [F], and then programming languages such as Plotkin's PCF can be translated into this, using the denotational semantics in [Plo77].

Computationally, we understand the elements $\top$ and $\bot$ of Σ as termination ("yes") and non-termination ("wait") respectively; Σ is not a "Boolean" type, as there is no *definite* "no". The naïve interpretation of the Phoa principle is then that any function $F : \Sigma \to \Sigma$ is specified by its effect on these two possible inputs. It cannot interchange them, as that would solve the Halting Problem, so $F\bot \Rightarrow F\top$, leaving essentially only three possible behaviours. More generally, all maps $\Sigma^Y \to \Sigma^X$ are monotone with respect to the lattice order in both topology and computation.

In fact, ASD is topologically more expressive than domain theory: its types denote exactly the "computably based" locally compact locales, and the maps are the computable continuous functions [G].

Conversely, the topological features of the calculus can be "normalised out" of its terms [A, §11]. Given that the most important type, Σ, captures *propositions* rather than *functions*, the result of this normalisation is a kind of *logic* program. However, to put this into practice would require a combination of *three* different generalisations of PROLOG, namely the addition of λ-calculus, parallelism (for $\vee$) and the manipulation of intervals or constraints.

In logic programming, a predicate is interpreted as a program whose objective is to find a proof and report instantiations for the free and existentially bound variables. If the predicate is false, the program never halts (in a "successful" way: it might abort). The basic process for doing this is unification, in which variables are assigned values according to the constraints imposed by the rest of the program, *i.e.* the values that they must have if the program is ever to terminate with a proof of the original predicate. The usual notion of unification only makes sense for combinatorial structures, but [Cle87] suggests an analogue for the reals.

Remark 3.5 Applying this to the issues of constructivity in the Introduction, logic programming gives an effective meaning to the *existential quantifier.* However, in ASD this is weaker than that found in other constructive foundational systems: in our version the free variables of the predicate must also be of overt discrete Hausdorff type. That is, they must be either natural numbers or something very similar, such as rationals. *They cannot be real numbers, functions or predicates.* Our ***choice principle*** is therefore only $\mathbb{N}$–$\mathbb{N}$ for Σ^0_1-predicates. This is topologically significant, because it constrains the sense in which the closed interval has the Heine–Borel "finite sub-cover" property in Section 10.

As you would expect, "de Morgan" duality between $\wedge$ and $\vee$ extends to the quantifiers, *i.e.* from $\exists$ to $\forall$. Once again, however, this is different in ASD from proof theory: whilst $\exists$ is true more often in ASD, $\forall$ is true less often, as we shall explain in Section 15.

4 The ASD lambda calculus

This section and the next summarise the symbolic language for ASD in a "user manual" style. Beware, however, that the manual for some device does not make its engineering redundant, and indeed the development of the more powerful devices of the future depends on advances in the underlying principles. Nevertheless, the constructions in the rest of this paper will be made entirely within this calculus. We repeat that there is no underlying set theory, and that the monadic and Phoa principles are unalienable parts of the system. On the other hand, there are other formulations of Scott continuity of various strengths besides the one that we give here.

Axiom 4.1 The ***types*** of ASD are generated from
(a) base types $\mathbf{1}$, Σ and $\mathbb{N}$, by
(b) products, $X \times Y$,
(c) exponentials, Σ^X, or $X \to \Sigma$ if you prefer, and
(d) Σ-split subspaces.

We do not introduce general exponentials Y^X, infinitary products, type variables or dependent types — at least, not in this version of the calculus. We will explain Σ-split subspaces, which abstract Proposition 2.15, in the next section; they will be used to define open, closed and retract subspaces, and, of course, R itself.

Axiom 4.2 The ***logical terms*** of type Σ or Σ^U (also known as ***propositions*** and ***predicates*** respectively) are generated from
(a) constants $\top$ and $\bot$;
(b) the lattice connectives $\wedge$ and $\vee$ (but not $\Rightarrow$ or $\neg$);
(c) λ-abstraction $\lambda x.\,\phi$, where ϕ must itself be logical (*i.e.* of some type of the form Σ^V, and in particular not $\mathbb{N}$ or $\mathbb{R}$), but x may be of any type;
(d) λ-application ϕa, where $a : A$, and $\phi : \Sigma^A$ is logical;
(e) equality $(n =_N m)$, where N is any *discrete* type;
(f) inequality or apartness $(h \neq_H k)$, where H is any *Hausdorff* type;
(g) existential quantification $\exists x{:}X.\,\phi x$, where ϕx is logical, and the type X is *overt*;
(h) universal quantification $\forall k{:}K.\,\phi k$, where ϕk is logical, and the type K is *compact*.
These terms are generated within a larger calculus that also includes
(i) variables of all types, for which we use increasingly exotic alphabets as the types get more complex; and
(j) pairing $\langle\,,\,\rangle$ and projections π_0, π_1 for product types.

Discrete, Hausdorff, overt and compact spaces will be defined shortly. Existential quantification over $\mathbb{N}$ or $\mathbb{R}$ is allowed, but universal quantification is not. Universal quantification over the *closed* interval $\mathbb{I} \equiv [0,1]$ is justified in Section 10.

So examples of *invalid* logical formulae include

$$\lambda n.\, n+3, \quad \lambda x.\, \sqrt[3]{x}, \quad \pi =_{\mathbb{R}} 3.14159, \quad \forall n.\, \exists pq.\, 2n = p+q \quad \text{and} \quad \forall x.\, x^2 \neq -1,$$

but $(\exists npqr{:}\mathbb{N}.\ p^n + q^n = r^n)$ and $(\forall x{:}|x| \leq 2.\ x^2 < 0)$ are fine. Programs to which an ML compiler would assign the types `nat` → `nat` or `real` → `real` are treated in ASD as terms of type $\mathbb{N} \to \mathbb{N}_\bot$ or $\mathbb{R} \to \mathbb{R}_\bot$ [D, F].

Binding, renaming, duplication, omission and substitution for variables are the same as in the λ- and predicate calculi. A quantified formula has the same type as its body (ϕ), whilst λ-abstraction and application modify types in the usual way. Alternating quantifiers are allowed to any depth — so long as their ranges are overt or compact spaces.

Axiom 4.3 The ***numerical terms*** of type $\mathbb{N}$ are generated from
(a) the constant 0;
(b) variables;
(c) successor; and
(d) definition by description: $\mathsf{the}\, n.\, \phi n$ (Axiom 4.24).
(e) We also allow *primitive recursion* over $\mathbb{N}$ at *any* type X.

Definition 4.4 ***Judgements*** in the calculus are of the four forms

$$\vdash\ X\ \mathsf{type}, \qquad \Gamma\ \vdash\ a : X, \qquad \Gamma\ \vdash\ a = b : X \quad \text{and} \quad \Gamma\ \vdash\ \alpha \leqslant \beta : \Sigma^X,$$

asserting well-formedness of types and typed terms, and equality or implication of terms. We shall refer to $a = b$ and $\alpha \leqslant \beta$ on the right of the $\vdash$ as ***statements***. On the left is the ***context*** Γ, consisting of assignments of types to variables, and maybe also equational hypotheses [E, §2].

The form $\Gamma\ \vdash\ \alpha \leqslant \beta : \Sigma^X$ is syntactic sugar:

Axiom 4.5 The predicates and terms satisfy certain ***equational axioms***, including
(a) those for a distributive lattice; in particular, for $\Gamma \vdash \phi, \psi : \Sigma^X$ we write

$$\phi\ \leqslant\ \psi \quad \text{to mean} \quad \phi \wedge \psi\ =\ \phi, \quad \text{or equivalently} \quad \phi \vee \psi\ =\ \psi,$$

although for $\sigma, \tau : \Sigma$ we use $\sigma \Rightarrow \tau$ and $\sigma \Leftrightarrow \tau$ instead of $\sigma \leqslant \tau$ and $\sigma = \tau$, since we shall also need the symbols $\leq$ and $=$ for their more usual "numerical" meanings;
(b) the β- and η-rules for λ-abstraction/application and pairing/projection;
(c) the β- and η-rules for primitive recursion over $\mathbb{N}$; and
(d) others that we describe in a little more detail in the rest of this section.

Remark 4.6 Predicates and terms on their own denote open subspaces and continuous functions respectively, but their expressive power is very weak. We introduce implications into the logic, but make their hierarchy explicit, in the form of statements and judgements.
(a) Observe, first, that $\leqslant$, $=$, $\Rightarrow$ and $\Leftrightarrow$ link *predicates* to form *statements*, not new predicates. In other words, Σ is a lattice and not a Heyting algebra.
(b) There are equality *statements* $a = b : X$ for *any* type, but equality *predicates* $(n =_N m) : \Sigma$ only for *discrete* types. In particular, for $a, b : R$, there are *predicates* $a \neq b$ and $a < b$, but $a = b$ and $a \leq b$ are *statements*, since R is Hausdorff but not discrete.
(c) For any proposition σ, we sometimes write σ or $\neg\sigma$ for the *statements* $\sigma \Leftrightarrow \top$ or $\sigma \Leftrightarrow \bot$.
(d) Nested equality and implication are not allowed in statements: we use a *judgement* of the form $\alpha_2 \Rightarrow \beta_1 \vdash \gamma_1 \Rightarrow \delta_0$ instead. If we need another level of nesting, we use a ***proof rule***:

$$\frac{\alpha_3 \Rightarrow \beta_2 \quad \vdash \quad \gamma_2 \Rightarrow \delta_1}{\epsilon_2 \Rightarrow \zeta_1 \quad \vdash \quad \eta_1 \Rightarrow \theta_0}$$

See Exercise 6.17 for some examples. This is as far as we can go. This rather artificial limit will be rectified in a future extended calculus [M].

(e) Equality and implication statements for function-types Σ^X provide a way of stating quantification over any type X — as a statement — it's a predicate only when X is compact. However, for the sake of clarity, we usually write $\Gamma, x : X \vdash \phi x \Leftrightarrow \psi x$ instead of $\Gamma \vdash \phi = \psi$.

Our convention is that λ-application binds most tightly, followed by the propositional relations, then $\wedge$, then $\vee$, then λ, $\exists$ and $\forall$, then $\Rightarrow$, $\Leftrightarrow$, $\leqslant$, $=$ and finally $\vdash$. This reflects the hierarchy of propositions, predicates, statements, judgements and rules. We often bracket propositional equality for emphasis and clarity.

Axiom 4.7 In addition to the equations for a distributive lattice, Σ satisfies the rules[1]

$$\frac{\Gamma,\ \sigma \Leftrightarrow \top\ \vdash\ \alpha\ \Rightarrow\ \beta}{\Gamma\ \vdash\ \sigma \wedge \alpha\ \Rightarrow\ \beta} \qquad\qquad \frac{\Gamma,\ \sigma \Leftrightarrow \bot\ \vdash\ \beta\ \Rightarrow\ \alpha}{\Gamma\ \vdash\ \beta\ \Rightarrow\ \sigma \vee \alpha}$$

Axiom 4.8 Every $F : \Sigma^\Sigma$ is monotone, *i.e.* $F\bot \Rightarrow F\top$.

Lemma 4.9 Σ satisfies the ***Phoa principle***, $F\sigma \iff F\bot \vee \sigma \wedge F\top$, for $F : \Sigma \to \Sigma$ and $\sigma : \Sigma$, possibly involving parameters. □

Plainly Phoa entails monotonicity, and we shall explain shortly why the Gentzen-style rules in Axiom 4.7 are also derivable from it.

Definition 4.10 We say that $\phi : \Sigma^X$ ***classifies*** an open subspace of X, and ***co-classifies*** a closed one. In symbols, for $a : X$,

$$a \in U \text{ (open) if } \phi a \Leftrightarrow \top \quad \text{and} \quad a \in C \text{ (closed) if } \phi a \Leftrightarrow \bot,$$

so $U \subset U'$ iff $C' \subset C$ iff $\phi \leqslant \phi'$. This means that

$$\text{open subspaces of } X, \quad \text{closed subspaces of } X \quad \text{and} \quad \text{continuous functions } X \to \Sigma$$

are in bijection. A ***clopen***, ***complemented*** or ***decidable*** subspace is one that is both open and closed. Then it and its complement are the inverse images of $0, 1 \in \mathbf{2}$ under some (unique) map $X \to \mathbf{2}$ [B, Theorem 11.8] [C, Proposition 9.6].

Remark 4.11 Such inverse image types cannot of course be generated from $\mathbb{N}$ and Σ by $\times$ and $\Sigma^{(-)}$. Our comments here will be justified by the introduction of open and closed subspaces as Σ-split subspaces in the next section. Briefly, being Σ-split means that all open subsubspaces of a subspace are restrictions of open subspaces of the ambient space, indeed in a canonical way.

Consider the left-hand Gentzen rule in terms of the open subspaces that α, β and σ classify. The top line expresses the containment $\alpha \Rightarrow \beta$ of open *sub*subspaces of the open subspace classified by σ. In the ambient space, this means that $\alpha \wedge \sigma \Rightarrow \beta \wedge \sigma$, or, more briefly, $\alpha \wedge \sigma \Rightarrow \beta$, as in the bottom line.

The rule on the right says exactly the same thing, but for the intersection of *closed* subspaces. Translating this into a result about relatively open subsubspaces of the closed subspace, the relative containment $\beta \Rightarrow \alpha$ in the closed subspace means $\beta \vee \sigma \Rightarrow \alpha \vee \sigma$, or just $\beta \Rightarrow \alpha \vee \sigma$, in the whole space.

[1] Peter Aczel pointed out the resemblance to Gerhard Gentzen's rules for negation in classical sequent calculus [Gen35].

Another "extensional" aspect of topology that the rules state is that, if the inverse images of $\top$ under $\phi, \psi : X \rightrightarrows \Sigma$ coincide as (open) subspaces of X, then $x : X \vdash \phi x \Leftrightarrow \psi x$ (or $\phi = \psi$) as logical terms. The inverse images of $\perp$ behave in a similar way [C, §5].

Subspaces of $X^{\mathbf{2}}$ or $X^{\mathbf{3}}$ are often called binary or ternary ***relations***. In particular, open binary relations are the same thing as predicates with two variables.

Definition 4.12 If the diagonal subspace, $X \subset X \times X$, is open or closed then we call X ***discrete*** or ***Hausdorff*** respectively. Type-theoretically, such spaces are those in which we may internalise equality-*statements* as *predicates*:

$$\frac{\Gamma \ \vdash \ n \ = \ m : N}{\Gamma \ \vdash \ (n =_N m) \ \Leftrightarrow \ \top} \qquad\qquad \frac{\Gamma \ \vdash \ h \ = \ k : H}{\Gamma \ \vdash \ (h \neq_H k) \ \Leftrightarrow \ \perp}$$

Lemma 4.13 Equality has the usual properties of substitution, reflexivity, symmetry and transitivity, whilst inequality or apartness obeys their lattice duals:

$$\begin{array}{rcl@{\qquad\qquad}rcl}
\phi m \wedge (n = m) & \Rightarrow & \phi n & \phi h \vee (h \neq k) & \Leftarrow & \phi k \\
(n = n) & \Leftrightarrow & \top & (h \neq h) & \Leftrightarrow & \perp \\
(n = m) & \Leftrightarrow & (m = n) & (h \neq k) & \Leftrightarrow & (k \neq h) \\
(n = m) \wedge (m = k) & \Rightarrow & (n = k) & (h \neq k) \vee (k \neq \ell) & \Leftarrow & (h \neq \ell)
\end{array}$$

(The proof of this uses Axiom 4.7.) In an overt discrete space, or a compact Hausdorff one, we have the converse of the first of the four rules:

$$\exists m.\, \phi m \wedge (n = m) \ \Leftrightarrow \ \phi n \qquad\qquad \forall h.\, \phi h \vee (h \neq k) \ \Leftrightarrow \ \phi k.$$

When *both* $=_X$ and $\neq_X$ are defined, as they are for $\mathbb{N}$, they are complementary:

$$(n =_X m) \vee (n \neq_X m) \ \Leftrightarrow \ \top \qquad \text{and} \qquad (n =_X m) \wedge (n \neq_X m) \ \Leftrightarrow \ \perp.$$

In this case X is said to have ***decidable equality***. $\square$

Definition 4.14 A space X that admits existential or universal quantification is called ***overt*** or ***compact*** respectively. By these quantifiers we mean the type-theoretic rules

$$\frac{\Gamma,\ x : X \ \vdash \ \phi x \ \Rightarrow \ \sigma}{\Gamma \ \vdash \ \exists x.\, \phi x \ \Rightarrow \ \sigma} \qquad\qquad \frac{\Gamma,\ x : X \ \vdash \ \sigma \ \Rightarrow \ \phi x}{\Gamma \ \vdash \ \sigma \ \Rightarrow \ \forall x.\, \phi x}$$

Axiom 4.15 The space $\mathbb{N}$ is overt. This and Scott continuity break the de Morgan-style lattice duality that the other rules enjoy.

Remark 4.16 So long as the types of the variables really are overt or compact, we may reason with the quantifiers in the usual ways:

(a) If we find a particular $a : X$ that satisfies ϕa, then we may of course assert $\exists x.\, \phi x$. This simple step tends to pass unnoticed in the middle of an argument, in the form $\phi a \Rightarrow \exists x.\, \phi x$.

(b) Similarly, if the judgement $\forall x.\, \phi x$ has been proved, and we have a particular value $a : X$, then we may deduce ϕa. Again, we often just write $(\forall x.\, \phi x) \Rightarrow \phi a$.

(c) The familiar mathematical idiom "there exists", in which $\exists x.\, \phi x$ is asserted and then x is temporarily used in the subsequent argument, is valid, as [Tay99, §1.6] explains.

(d) The λ-calculus formulation automatically allows substitution under the quantifiers [C, §8], whereas in categorical logic this property must be stated separately, and is known as the Beck–Chevalley condition [Tay99, Chapter IX].

However, we must observe the caveats in Remarks 3.5 and 15.4:

(e) Given $\exists x.\, \phi x$, there need only be a term $a : X$ satisfying ϕa in the case where X and the types all of the parameters in ϕ are overt discrete and Hausdorff.

(f) We cannot deduce $\forall x.\, \phi x$ from proofs of ϕa for every closed term $a : X$.

Exercise 4.17 Use the Phoa principle to prove the ***Frobenius*** and ***modal laws***

$$\begin{array}{lcl@{\qquad}lcl} \exists x.\, \sigma \wedge \phi x & \Leftrightarrow & \sigma \wedge \exists x.\, \phi x & (\forall x.\, \phi x) \wedge (\exists x.\, \psi x) & \Rightarrow & \exists x.\, (\phi x \wedge \psi x) \\ \forall x.\, \sigma \vee \phi x & \Leftrightarrow & \sigma \vee \forall x.\, \phi x & (\forall x.\, \phi x) \vee (\exists x.\, \psi x) & \Leftarrow & \forall x.\, (\phi x \vee \psi x) \end{array}$$

where the type of x is both overt and compact. The Frobenius law for $\forall$ is another feature that ASD has in common with classical but not intuitionistic logic; it was nevertheless identified in intuitionistic locale theory by Japie Vermeulen [Ver94]. $\square$

Lemma 4.18 Any topology Σ^X has joins indexed by overt objects and meets indexed by compact ones:

$$\bigvee_N \equiv \exists^X_N : (\Sigma^X)^N \cong (\Sigma^N)^X \longrightarrow \Sigma^X \quad \text{and} \quad \bigwedge_K \equiv \forall^X_K : (\Sigma^X)^K \cong (\Sigma^K)^X \longrightarrow \Sigma^X.$$

Binary meets distribute over joins by the Frobenius law, and "substitution under the quantifier" means that all inverse image maps $\Sigma^f : \Sigma^Y \to \Sigma^X$, where

$$\Sigma^f \psi \;\equiv\; \psi \cdot f \;\equiv\; \lambda x.\, \psi(fx) \quad \text{or} \quad \Sigma^f V \;\equiv\; f^* V \;\equiv\; f^{-1} V \;\equiv\; \{x \mid \psi x\},$$

preserve all joins indexed by overt objects, and meets indexed by compact ones [C, §7]. $\square$

Since $\mathbb{N}$ is overt but not compact, each Σ^X has and each Σ^f preserves $\mathbb{N}$-indexed joins, but not $\mathbb{N}$-indexed meets.

Remark 4.19 We often want the quantifiers or meets and joins to range over dependent types, even though we have not provided these in the calculus.

The most pressing case of this is the join or existential quantifier indexed by an open subspace $M \subset N$ of an overt space. This subspace is classified by a predicate $\alpha : \Sigma^N$, which we shall write as $\Gamma,\ n : N\ \vdash\ \alpha_n : \Sigma$. The M-indexed family $\phi^m : \Sigma^X$ of which we want to form the join may always be considered to be the restriction to M of an N-indexed family, so we have

$$\bigvee_{m:M} \phi^m \;\equiv\; \exists n{:}N.\, \alpha_n \wedge \phi^n : \Sigma^X.$$

This sub- and super-script notation, which is used extensively in [G], indicates co- and contravariance with respect to an imposed order relation, such as the arithmetical order in $\mathbb{Q}$ or inclusion of lists.

We shall also want to define both quantifiers over the closed interval $[d, u]$, where d and u need not be constants, but we shall cross this bridge when we come to it, in Section 10.

Definition 4.20 Such a pair of families (α_n, ϕ^n) is called a ***directed diagram***, and the corresponding

$$\bigvee\!\!\!\!\nearrow_{n:\alpha_n} \phi^n \quad \equiv \quad \exists n.\, \alpha_n \wedge \phi^n$$

is called a ***directed join***, if (a) $(\exists n.\, \alpha_n) \Leftrightarrow \top$, and (b) $\alpha_{n@m} \Leftrightarrow \alpha_n \wedge \alpha_m$ and $\phi^{n@m} \geqslant \phi^n \vee \phi^m$ for some binary operation $@ : N \times N \to N$.

In this, $\alpha_n \wedge \alpha_m$ means that both ϕ^n and ϕ^m contribute to the join, so for directedness in the informal sense, we require some $\phi^{n@m}$ to be above them both (contravariance), and also to count towards the join, for which $\alpha_{n@m}$ must be true (covariance). Hence the imposed order relation on N is that for which $(N, @)$ is essentially a meet semilattice.

This allows us to formalise ***Scott continuity*** (Definitions 2.6 and 3.1).

Axiom 4.21 Any $F : \Sigma^{\Sigma^X}$ (possibly with parameters) preserves directed joins in the sense that

$$F(\exists n.\, \alpha_n \wedge \phi^n) \iff \exists n.\, \alpha_n \wedge F\phi^n.$$

Notice that F is attached to ϕ^n and not to α_n, since the join being considered is really that over the subset $M \equiv \{n \mid \alpha_n\} \subset N$. The principal use of this axiom in this paper is Proposition 7.12, where the ambient overt object N is $\mathbb{Q}$, and its imposed order is either the arithmetic one or its reverse, so @ is either max or min. Scott continuity is also what connects our definition of compactness with the traditional "finite open sub-cover" one (Remark 10.10), whilst in denotational semantics it gives a meaning to recursive programs.

In locale theory [Joh82], the homomorphisms are those of finite meets and arbitrary joins. Since we have just taken care of (some of) the *directed* joins, in ASD we need only consider the finitary connectives $\top$, $\bot$, $\wedge$ and $\vee$.

Definition 4.22 A term P of type Σ^{Σ^X} is called ***prime*** if it preserves all four lattice connectives:

$$P\top \Leftrightarrow \top \qquad P\bot \Leftrightarrow \bot \qquad P(\phi \wedge \psi) \Leftrightarrow P\phi \wedge P\psi \qquad P(\phi \vee \psi) \Leftrightarrow P\phi \vee P\psi.$$

The type X is said to be ***sober*** if every prime P corresponds to a unique term $a : X$ (which we call $\mathsf{focus}\, P$) such that

$$\phi(\mathsf{focus}\, P) \Leftrightarrow P\phi \qquad \text{or} \qquad P \;=\; \lambda\phi.\, \phi a.$$

Proposition 4.23 Let $i : X \to Y$ between sober objects. If the map $\Sigma^i : \Sigma^Y \to \Sigma^X$ has an inverse I, then so does i itself [Tay91].

Proof The map $\eta : Y \rightarrowtail \Sigma^{\Sigma^Y}$ given by $y \mapsto \lambda\psi.\, \psi y$ is mono, because $\mathsf{focus}_Y(\lambda\psi.\, \psi y) = y$.

Since I is an isomorphism, both it and $P \equiv \lambda\phi.\, I\phi y$ preserve all four lattice operations, *i.e.* P is prime. Then, as X is sober, we may define $j : Y \to X$ by $jy \equiv \mathsf{focus}_X(\lambda\phi.\, I\phi y)$. It satisfies

$$j(ix) \;=\; \mathsf{focus}_X\big(\lambda\phi.\, I\phi(ix)\big) \;=\; \mathsf{focus}_X\big(\lambda\phi.\, I(\Sigma^i\phi)x\big) \;=\; \mathsf{focus}_X(\lambda\phi.\, \phi x) \;=\; x$$

$$\text{and} \quad \lambda\psi.\, \psi\big(i(jy)\big) \;=\; \lambda\psi.\, (\Sigma^i\psi)\big(\mathsf{focus}_X(\lambda\phi.\, I\phi y)\big) \;=\; \lambda\psi.\, I(\Sigma^i\psi)y \;=\; \lambda\psi.\, \psi y \;\equiv\; \eta_Y y,$$

so $i(jy) = y$ since $\eta : Y \rightarrowtail \Sigma^{\Sigma^Y}$ is mono. $\square$

For $\mathbb{N}$, sobriety is equivalent to a more familiar idiom of reasoning that characterises *singleton* predicates, $\phi = \{n\} \equiv (\lambda m.\, n = m)$. This is only meaningful for (overt) *discrete* objects, but we shall see in Section 14 that Dedekind completeness expresses sobriety of $\mathbb{R}$.

Axiom 4.24 A predicate ϕn for $n : \mathbb{N}$ (possibly involving other parameters) is called a ***description*** if it is "uniquely satisfiable" in the sense that

$$(\exists n.\, \phi n) \iff \top \quad \text{and} \quad (\phi n \wedge \phi m) \implies (n =_{\mathbb{N}} m)$$

are provable. Then we may introduce $\mathsf{the}\; n.\, \phi n : \mathbb{N}$, with the same parameters, satisfying

$$\phi m \iff (m =_{\mathbb{N}} \mathsf{the}\; n.\, \phi n).$$

Descriptions may be used to define general recursion (minimalisation) [A, D, O].

Proposition 4.25 All types that are definable from $\mathbf{1}$, Σ and $\mathbb{N}$ using $\times$ and $\Sigma^{(-)}$ are sober.

Proof

(a) The only prime of type $\Sigma^{\Sigma^{\mathbf{1}}}$ is $P \equiv \lambda\phi.\ \phi\star$, and $\mathsf{focus}_{\mathbf{1}}\, P = \star$.

(b) If $P : \Sigma^{\Sigma^{\mathbb{N}}}$ is prime then $\phi n \equiv P(\lambda m.\ n =_{\mathbb{N}} m)$ is a description, and $\mathsf{focus}_{\mathbb{N}}\, P = \mathsf{the}\ n.\ \phi n$ [A, Proposition 10.4].

(c) If X and Y are sober then so is $X \times Y$, with

$$\mathsf{focus}_{X\times Y}\, P \;=\; \left\langle \mathsf{focus}_X \left(\lambda\phi.\ P(\lambda p.\ \phi(\pi_0 p))\right), \mathsf{focus}_Y \left(\lambda\psi.\ P(\lambda p.\ \psi(\pi_1 p))\right)\right\rangle.$$

(d) Any Σ^X is sober, with $\mathsf{focus}_{\Sigma^X}\, \mathcal{P} \equiv \lambda x.\ \mathcal{P}(\lambda\phi.\ \phi x)$ [A, Lemma 8.8]. $\square$

Examples 4.26 We conclude with some examples of the four main topological properties.

	overt	discrete	compact	Hausdorff
$\mathbb{N} \times \Sigma$	✓	×	×	×
$\mathbb{R}$, $\mathbb{R}^{\mathbf{n}}$	✓	×	×	✓
Σ	✓	×	✓	×
$\mathbb{I}$, $\mathbf{2}^{\mathbb{N}}$	✓	×	✓	✓
free combinatory algebra	✓	✓	×	×
$\mathbb{N}$, $\mathbb{Q}$	✓	✓	×	✓
Kuratowski finite	✓	✓	✓	×
finite ($\mathbf{n}$)	✓	✓	✓	✓
the set of codes of non-terminating programs	×	✓	×	✓

The axioms only provide the quantifiers $\forall_\emptyset \equiv \top$, $\exists_\emptyset \equiv \bot$, $\forall_{\mathbf{2}} \equiv \wedge$, $\exists_{\mathbf{2}} \equiv \vee$ and $\exists_{\mathbb{N}}$ directly — it is the business of this paper to *construct* $\exists_{\mathbb{R}}$, $\exists_{\mathbb{I}}$ and $\forall_{\mathbb{I}}$, in Sections 9 and 10. To put this the other way round, assuming that a future extension of the calculus will allow arbitrary nesting of $\Rightarrow$ and $\forall$ in statements (*cf.* Remark 4.6), these constructions will justify ***quantifier elimination*** in the corresponding cases.

The calculus as we have described it so far will be used to define *individual* Dedekind cuts in Section 6.

5 The monadic principle

We have seen how some of the ideas of topology can be expressed using the types and terms of a λ-calculus. In particular, compactness of a space K is expressed as a "universal quantifier" $\forall_K : \Sigma^K \to \Sigma$. The main application of this idea will be to the closed interval $K \equiv [d,u] \subset R \subset \Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}} \equiv Y$, where we shall construct R as the *subspace* of those pairs of predicates on $\mathbb{Q}$ that are Dedekind cuts (Remark 2.1).

How could we obtain a map $\Sigma^K \to \Sigma$ from one $i : K \rightarrowtail Y$? We have $\Sigma^Y \to \Sigma$ by evaluation at some $y : Y$ (Lemma 5.12), but nothing $\Sigma^K \to \Sigma^Y$, since the maps go the wrong way. However, given that we have already introduced higher *types*, it hardly seems to be a revolutionary step to introduce extra *terms* of those types. Indeed, in our brief account of interval analysis in Section 2, we have already discussed such a map, $I : \Sigma^K \to \Sigma^Y$. It extends open sets and functions from the real line to the interval domain, and it is intimately related to the quantifiers $\forall$ and $\exists$, indeed these are special cases of it.

In this section we shall give the type-theoretic rules that formally introduce i and I for (certain) subspaces, at least the ones that we need. This will complete our manual of the axioms of ASD.

The basis of our proof of the Heine–Borel theorem is that the rules force these subspaces to carry the subspace topology. Naturally, we shall use terms of higher type (such as I) to do this, just as we have already introduced terms of lower type and equations between them in the previous section.

Definition 5.1 An inclusion $i : X \rightarrowtail Y$ is called a Σ-***split subspace*** if there is another map $I : \Sigma^X \rightarrowtail \Sigma^Y$ such that

$$\Sigma^i \cdot I \;=\; \mathsf{id}_{\Sigma^X} \qquad \text{or} \qquad x : X,\ \phi : \Sigma^X \;\vdash\; I\phi(ix) \;\Leftrightarrow\; \phi x,$$

thereby making Σ^X a retract of Σ^Y with inclusion I and projection Σ^i. The convention in the λ-calculus is that $I\phi(ix)$ means $(I\phi)(ix)$.

Remark 5.2 Whenever two maps compose to the identity one way round like this, they may be recovered (uniquely up to unique isomorphism) from their composite the other way, which is an idempotent ($E = E \cdot E$). In this case, $E \equiv I \cdot \Sigma^i : \Sigma^Y \to \Sigma^Y$ determines Σ^X. (We wrote $\mathcal{E}$ for E in Proposition 2.15 in accordance with our convention of using fancier typefaces for fancier types.) The object X itself is the equaliser

$$\begin{tikzcd} X \arrow[r, tail, "i"] & Y \arrow[r, shift left, "y \mapsto \lambda\psi.\, \psi y"] \arrow[r, shift right, "y \mapsto \lambda\psi.\, E\psi y"'] & \Sigma^{\Sigma^Y}. \end{tikzcd}$$

The idea is that the operator E "normalises" open subspaces of Y by restricting them to X using Σ^i, and then re-expanding them using I. Therefore $a : Y$ belongs to $\{Y \mid E\}$ iff its membership of *any* open subspace ψ of Y is unaffected by this process.

Definition 5.3 Recall that being the ***equaliser*** means that
(a) the inclusion i has equal composites with the parallel pair: it sends each $x : X$ to $y \equiv ix : Y$ satisfying the equation ($\psi y \Leftrightarrow E\psi y$ for all $\psi : \Sigma^Y$);
(b) whenever we have another map $b : \Gamma \to Y$ (which is the same as a term $b : Y$ with parameters from Γ) that has equal composites with the parallel pair (*i.e.* which satisfies the equation $\psi b \Leftrightarrow E\psi b$ for all ψ), then we may introduce a map $a : \Gamma \to X$ (term $\Gamma \vdash a : X$) that makes the triangles commute, so $b = ia$, and this is unique.

Not every idempotent E on Σ^Y can arise in this way from a Σ-split subspace, because the surjective part Σ^i of its splitting must be a homomorphism for the topological structure. In particular, it must preserve the four lattice connectives, but any order-preserving *surjective* function between lattices preserves $\top$ and $\bot$ anyway. So the crucial condition is this:

Lemma 5.4 $E \equiv I \cdot \Sigma^i : \Sigma^Y \to \Sigma^Y$ satisfies the equations, for $\phi, \psi : \Sigma^Y$,

$$E(\phi \wedge \psi) \;=\; E(E\phi \wedge E\psi) \quad \text{and} \quad E(\phi \vee \psi) \;=\; E(E\phi \vee E\psi). \qquad \square$$

Definition 5.5 A term E that satisfies the equations in the Lemma is called a ***nucleus***. In order to avoid developing a theory of dependent types, we do not allow E to have parameters.

The word "nucleus" was appropriated from locale theory, since both kinds of nuclei play the same role, namely to define subspaces, but the definitions are different. A localic nucleus, usually called j, must satisfy $\mathsf{id} \leqslant j = j \cdot j$ but need not be Scott-continuous [Joh82, §II 2]. Nuclei in ASD, on the other hand, are continuous but need not be order-related to id. So the common ground is when $E \geqslant \mathsf{id}$ and j is continuous: ASD and locale theory therefore agree on how to form closed subspaces, but not open ones or retracts.

In Section 8 we shall use $\mathcal{E}$ to define R as a subspace of the interval domain IR, and in this case $\mathcal{E}$ is a nucleus in both senses. If we view R instead as a subspace of $\Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}$, but still defined by $\mathcal{E}$, this is a nucleus only in the sense of ASD, not that of locale theory.

Notation 5.6 We write $\{Y \mid E\}$ for the subspace $X \subset Y$ that is defined by any nucleus E on Y. Beware, however, that this has very few of the properties of subset-formation in set theory.

Definition 5.7 A term $a : Y$ (which may now involve parameters from a context Γ) of the larger space Y is called ***admissible*** with respect to the nucleus E if

$$\Gamma,\ \psi : \Sigma^Y \ \vdash\ \psi a \ \Leftrightarrow\ E\psi a.$$

This is the same as saying that the map $a : \Gamma \to Y$ has equal composites with the parallel pair in the diagram in Remark 5.2, so it factors through the equaliser.

Since this equaliser is the subtype $X \subset Y$, we may then regard a as a term of type X or $\{Y \mid E\}$. However, when we need to discuss the calculus in a formal way, we shall write $\mathsf{admit}_{Y,E}\, a$ for it, in order to disambiguate its type.

Remark 5.8 The simplest (and first, [C, §3]) application of this calculus is to the construction of open and closed subspaces. This will formalise Remark 4.11.

Let $\theta : \Sigma^Y$ be a closed term (*i.e.* without parameters), which we think of as a continuous function $\theta : Y \to \Sigma$. The open and closed subspaces that θ (co)classifies are the inverse images $U \equiv \theta^{-1}(\top)$ and $C \equiv \theta^{-1}(\bot)$ of the two points of the Sierpiński space (Remark 3.2).

Classically, any open subspace of U is already one of Y, whilst an open subspace of C becomes open in Y when we add U to it. This means that

$$\Sigma^U \ \cong\ \Sigma^Y \downarrow \theta \ \equiv\ \{\psi \mid \psi \leqslant \theta\} \quad \text{and} \quad \Sigma^C \ \cong\ \theta \downarrow \Sigma^Y \ \equiv\ \{\psi \mid \psi \geqslant \theta\},$$

i.e. the parts of the lattice Σ^Y that lie respectively below and above θ. (The $\downarrow$ notation is used in category theory for "slice" or "comma" categories.)

From this we see what the Σ-splittings and nuclei for U and C must be. The construction of the interval domain $IR \subset \Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}$ in Section 7 will involve an open subspace, a closed one and a retract.

Lemma 5.9 Let $e : Y \to Y$ be an idempotent and $\theta : Y \to \Sigma$ any map. Then

$$\Sigma^e, \qquad \theta \wedge (-) \qquad \text{and} \qquad \theta \vee (-)$$

are nuclei on Y, with respect to which $a : Y$ is admissible iff respectively

$$ea = a, \qquad \theta a \Leftrightarrow \top \qquad \text{and} \qquad \theta a \Leftrightarrow \bot. \qquad \square$$

Notation 5.10 As we shall need to name the Σ-splittings of $j : U, C \rightarrowtail Y$, we write

$$\exists_j : \Sigma^U \rightarrowtail \Sigma^Y \qquad \text{and} \qquad \forall_j : \Sigma^C \rightarrowtail \Sigma^Y,$$

where $\exists_j(\Sigma^j\psi) = \psi \wedge \theta$ and $\forall_j(\Sigma^j\psi) = \psi \vee \theta$. See [C] for the reasons for these names.

In Sections 7 and 8 we shall need to combine several nuclei. However, the intersection of two Σ-split subspaces need not exist, in general, and even if it does, it need not be Σ-split. Everything works if the nuclei commute, but this does not often happen. In fact, there are many situations in which $E_1 \cdot E_2 \cdot E_1 = E_1 \cdot E_2$, and that is sufficient to justify forming the intersection.

Lemma 5.11 Let E_1 and E_2 be nuclei on a type Y such that either $E_1 \cdot E_2 = E_2 \cdot E_1$ or $E_1 \cdot E_2 \cdot E_1 = E_1 \cdot E_2$. Then $E_1 \cdot E_2$ is also a nucleus, and defines the intersection (pullback)

$$\begin{tikzcd} \{Y \mid E_1 \cdot E_2\} \arrow[r, tail] \arrow[d, tail] \arrow[dr, phantom, "\lrcorner", very near start] & \{Y \mid E_1\} \arrow[d, tail] \\ \{Y \mid E_2\} \arrow[r, tail] & Y \end{tikzcd}$$

in which all five inclusions (including the composite) are Σ-split. Regarding $\{Y \mid E_1\} \subset Y$ as the subset of admissible elements, and $\Sigma^{\{Y|E_1\}}$ as the subset $\{\psi : \Sigma^Y \mid E_1\psi = \psi\}$, we see $E_1 \cdot E_2$ as a nucleus on $\{Y \mid E_1\}$ but not necessarily on $\{Y \mid E_2\}$. □

We have already said that nuclei and Σ-splittings are intimately related to the quantifiers that define compactness and overtness. The following two lemmas will provide the key to defining $\exists$ and $\forall$ from I in Section 10. There we shall use the slices $\Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}} \downarrow (\delta_u, \upsilon_d)$ and $(\delta_d, \upsilon_u) \downarrow \Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}$, which we shall call ***lattices*** in this very specific context.

Lemma 5.12 Any lattice L is overt and compact, with $\exists\Phi \equiv \Phi(\top)$ and $\forall\Phi \equiv \Phi(\bot)$.

Proof By monotonicity, $\Phi(\bot) \Rightarrow \Phi(x) \Rightarrow \Phi(\top)$ for any $x : L$. □

Lemma 5.13 Let $i : X \rightarrowtail L$ and $I : \Sigma^X \rightarrowtail \Sigma^L$ with $\Sigma^i \cdot I = \mathsf{id}$. Then[2]

(a) if I preserves $\bot$ and L is overt then so is X, with $\exists x{:}X.\ \phi x \equiv \exists y{:}L.\ I\phi y$, whilst

(b) if I preserves $\top$ and L is compact then so is X, with $\forall x{:}X.\ \phi x \equiv \forall y{:}L.\ I\phi y$.

Proof In each column we deduce in both directions,

$$\begin{array}{lll} \Gamma,\ x : X & \vdash & \phi x \ \Rightarrow\ \sigma \\ \Gamma,\ \sigma \Leftrightarrow \bot & \vdash & \Sigma^i(I\phi)\ \equiv\ \phi\ =\ \bot : \Sigma^X \\ \Gamma,\ \sigma \Leftrightarrow \bot & \vdash & I\phi\ =\ \bot : \Sigma^L \\ \Gamma,\ \sigma \Leftrightarrow \bot & \vdash & \exists_L(I\phi)\ \Rightarrow\ \bot \\ \Gamma & \vdash & \exists_L(I\phi)\ \Rightarrow\ \sigma \end{array} \qquad \begin{array}{lll} \Gamma,\ x : X & \vdash & \sigma\ \Rightarrow\ \phi x \\ \Gamma,\ \sigma \Leftrightarrow \top & \vdash & \top\ =\ \phi\ \equiv\ \Sigma^i(I\phi) : \Sigma^X \\ \Gamma,\ \sigma \Leftrightarrow \top & \vdash & \top\ =\ I\phi : \Sigma^L \\ \Gamma,\ \sigma \Leftrightarrow \top & \vdash & \top\ \Rightarrow\ \forall_L(I\phi) \\ \Gamma & \vdash & \sigma\ \Rightarrow\ \forall_L(I\phi) \end{array}$$

using Axiom 4.7 and Definition 4.14. □

It is high time we gave the actual rules that define subtypes and their admissible terms. These *formally adjoin* Σ-split subspaces to the calculus just as number theorists formally adjoin roots of particular polynomials to a *field*, or set theorists construct "equiconsistent" *models* that have altered properties. Theorem 15.8 describes the corresponding categorical construction.

As in type theory, we name the rules after the connective that they introduce or eliminate. Our connectives are called $\{\}$ and $\Sigma^{\{\}}$.

Axiom 5.14 The $\{\}$-rules of the ***monadic*** $\boldsymbol{\lambda}$***-calculus*** define the subspace itself:

$$\frac{Y\ \mathsf{type} \qquad y : Y,\ \psi : \Sigma^Y\ \vdash\ E\psi y : \Sigma \qquad E \text{ is a nucleus}}{\{Y \mid E\}\ \mathsf{type}} \qquad \{\}F$$

[2]This result was suggested by Alex Simpson.

$$\frac{\Gamma \vdash a : Y \qquad \Gamma,\ \psi : \Sigma^Y \vdash \psi a \Leftrightarrow E\psi a}{\Gamma \vdash \mathsf{admit}_{Y,E}\, a : \{Y \mid E\}} \qquad \{\}I$$

$$x : \{Y \mid E\} \vdash i_{Y,E}x : Y \qquad \{\}E_0$$

$$x : \{Y \mid E\},\ \psi : \Sigma^Y \vdash \psi(i_{Y,E}x) \Leftrightarrow E\psi(i_{Y,E}x) \qquad \{\}E_1$$

$$\frac{\Gamma \vdash a : Y \qquad \Gamma,\ \psi : \Sigma^Y \vdash \psi a \Leftrightarrow E\psi a}{\Gamma \vdash a = i_{Y,E}(\mathsf{admit}_{Y,E}\, a) : Y} \qquad \{\}\beta$$

$$x : X \equiv \{Y \mid E\} \vdash x = \mathsf{admit}_{Y,E}(i_{Y,E}x) : \{Y \mid E\}. \qquad \{\}\eta$$

These rules say that $\{Y \mid E\}$ is the equaliser in Remark 5.2.

Axiom 5.15 The following $\Sigma^{\{\}}$-rules say that $\{Y \mid E\}$ has the subspace topology, where $I_{Y,E}$ expands open subsets of the subspace to the whole space.

$$\phi : \Sigma^{\{Y|E\}} \vdash I_{Y,E}\phi : \Sigma^Y. \qquad \Sigma^{\{\}}E$$

The $\Sigma^{\{\}}I$ rule is derived from $\{\}E_0$:

$$\psi : \Sigma^Y \vdash \Sigma^i\psi \equiv \lambda x{:}\{Y \mid E\}.\ \psi(i_{Y,E}x) : \Sigma^{\{Y|E\}}. \qquad \Sigma^{\{\}}I$$

The β-rule says that the composite $\Sigma^Y \xrightarrow{\Sigma^i} \Sigma^{\{Y|E\}} \xrightarrow{I} \Sigma^Y$ is E:

$$y : Y,\ \psi : \Sigma^Y \vdash I_{Y,E}\big(\lambda x{:}\{Y \mid E\}.\ \psi(i_{Y,E}x)\big)y \Leftrightarrow E\psi y. \qquad \Sigma^{\{\}}\beta$$

The η-rule says that the other composite $\Sigma^{\{Y|E\}} \longrightarrow \Sigma^Y \longrightarrow \Sigma^{\{Y|E\}}$ is the identity:

$$\phi : \Sigma^{\{Y|E\}},\ x : \{Y \mid E\} \vdash \phi x \Leftrightarrow I_{Y,E}\phi(i_{Y,E}x). \qquad \Sigma^{\{\}}\eta$$

This equation is the one in Definition 5.1 for a Σ-split subspace.

Remark 5.16 Notice that the $\Sigma^{\{\}}\beta$-rule is the only one that introduces E into terms. This is something that we want to avoid if at all possible, given that the expressions for nuclei, such as $\mathcal{E}$, which we intend to use to construct R, are usually very complicated. The normalisation theorem that eliminates unnecessary Es is proved in [B, §§9-10].

When E does find its way into a program, it will give rise to a substantial computation. But this is what we would expect, considering the relationship to the quantifiers, and their computational meaning as optimisation and search (Section 2).

Warning 5.17 It is important to understand that, for another structure to admit a valid interpretation of this calculus, not only must the equaliser in Remark 5.2 exist, but there must also be a map I in the structure that satisfies the equations $\Sigma^i \cdot I = \mathsf{id}_{\Sigma^X}$ and $I \cdot \Sigma^i = E$, *cf.* Remark 2.18.

Proposition 5.18
(a) If Y is sober then so is $\{Y \mid E\}$, with $\mathsf{focus}_{\{Y|E\}}\, Q = \mathsf{admit}_{Y,E}\big(\mathsf{focus}_Y(\Sigma^2 i_{Y,E} Q)\big)$.
(b) If X is also sober and has maps $f : X \to Y$ and $F : \Sigma^X \to \Sigma^Y$ with $\Sigma^f \cdot F = \mathsf{id}_{\Sigma^X}$ and $F \cdot \Sigma^f = E$ then $X \cong \{Y \mid E\}$, where

$$\begin{array}{lll} x : X & \vdash & x' \equiv \mathsf{admit}_{Y,E}(fx) : \{Y \mid E\} \\ x' : \{Y \mid E\} & \vdash & x \equiv \mathsf{focus}_X\big(\lambda\phi{:}\Sigma^X.\ F\phi(ix')\big) : X. \end{array}$$

In particular, f is the equaliser in Remark 5.2. $\square$

Remark 5.19 If this paper is the first one about ASD that you have seen, you may reasonably imagine that we have cooked up the rules from the observations in Section 2, based on little more than the properties of the real line that first year undergraduates learn. Can this theory of topology even be generalised to $\mathbb{R}^{\mathbf{n}}$? Is it of any use in proving theorems in mathematics other than those that we have mentioned?

Actually, we have told the story so far in reverse historical order. The calculus above was derived from a categorical intuition, and when it was formulated [B], it was a solution in search of a problem. There were no clear ideas of how the real line might be constructed, and certainly no inkling that it might behave in ASD in a radically different way from the established theory of Recursive Analysis (Section 15). That it turned out to do so is a tribute to category theory as a source of, or rather as a way of communicating, mathematical intuitions.

For reference we give a summary of some of the constructions that have been performed using nuclei in the infrastructure of ASD.

Theorem 5.20 Σ-split subspaces interact with the underlying structure of products and exponentials as follows, and also provide stable disjoint unions [B, §11].

$$\begin{array}{lcl}
X & \cong & \{\Sigma^{\Sigma^X} \mid \lambda\mathcal{F}F.\, F\big(\lambda x.\, \mathcal{F}(\lambda\phi.\, \phi x)\big)\} \\
\{X \mid E_0\} \times \{Y \mid E_1\} & \cong & \{X \times Y \mid E_1^X \cdot E_0^Y\} \\
\Sigma^{\{X|E\}} & \cong & \{\Sigma^X \mid \Sigma^E\} \\
\{\{X \mid E_1\} \mid E_2\} & \cong & \{X \mid E_2\} \\
\{X \mid E_0\} + \{Y \mid E_1\} & \cong & \{\Sigma^{\Sigma^X \times \Sigma^Y} \mid \mathsf{E}\} \\
\text{where} \quad \mathsf{E}\mathcal{H}H & \equiv & H\big\langle \lambda x.\, \mathcal{H}(\lambda\phi\psi.\, E_0\phi x), \lambda y.\, \mathcal{H}(\lambda\phi\psi.\, E_1\psi y)\big\rangle.
\end{array}$$
$\square$

In Section 8, we shall need to perform some combinatorial arguments using lists of rational numbers. However, since ASD is a direct axiomatisation of topology, without using set theory, we have to recover some kind of "naïve set theory" within our category of spaces. Our "sets" are the overt discrete objects.

Theorem 5.21 The full subcategory consisting of the overt discrete objects has
(a) finite products ($\mathbf{1}$ and $\times$);
(b) equalisers;
(c) stable disjoint coproducts ($\mathbf{0}$ and $+$);
(d) stable effective quotients of open equivalence relations [C]; and
(e) free monoids or "sets of lists", $\mathsf{List}(X)$ [E].
This categorical structure is called an ***arithmetic universe***. $\square$

Remark 5.22 We shall rely on this result to provide the "discrete mathematics" that we need, namely
(a) the usual arithmetic operations on $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$;
(b) the operations of *naïve* set theory over these types, including, for example, the use of lists and (Kuratowski-) finite sets; and
(c) both quantifiers over finite sets, *i.e.* finite conjunctions and disjunctions, which have well known definitions using primitive recursion.

The infinite objects that are generated in this way are ***countable*** in the strict sense of being bijective with a decidable subset of $\mathbb{N}$ in the calculus; since our isomorphisms are homeomorphisms, such objects inherit $=$, $\exists$, and in some cases $\neq$, from $\mathbb{N}$.

These "sets" do not admit a full powerset on the basis of the axioms that we have given here, and therefore they do not form an elementary topos or model of set theory. However, it is possible to add another (much more powerful, and certainly non-computational) axiom to ASD that does make the full subcategory of overt discrete objects into a topos. This axiom asserts that this subcategory is co-reflective, *i.e.* that the inclusion has a right adjoint, which formally assigns an overt discrete object to any space, and we call this object the "underlying set" of the space [H]. We do not use this axiom in this paper.

Remark 5.23 In order to formulate the definition of Σ-split subspace we only need powers $\Sigma^{(-)}$, not the lattice structure on Σ or the quantifiers. The category-theoretic summary of this section is that we require the adjunction $\Sigma^{(-)} \dashv \Sigma^{(-)}$ to be monadic.

The notions of sobriety and Σ-split subspaces in ASD come from the two conditions in the theorem of Jon Beck (1966) that characterise monadic adjunctions [Tay99, Theorem 7.5.9]. Sobriety says that $\Sigma^{(-)}$ reflects invertibility (Proposition 4.23). Nuclei are essentially the same as $\Sigma^{(-)}$-contractible equalisers; Axiom 5.14 says that such equalisers exist, and Axiom 5.15 that $\Sigma^{(-)}$ takes them to coequalisers. These conditions are given as equations between higher-order λ-terms in [A, B].

In the presence of the other topological structure in the previous section, in particular Scott continuity, the equations that define a nucleus are equivalent to the much simpler ones involving finitary $\wedge$ and $\vee$ that we gave in Lemma 5.4 and Definition 4.22 [G, §10]. In this setting, any definable object X of ASD can be expressed as a Σ-split subspace, $i : X \rightarrowtail \Sigma^{\mathbb{N}}$, where i and I correspond directly to a basis of open subspaces and related compact subspaces. Locally compact spaces and computably continuous functions between them can then be understood directly in ASD [G]. This calculus is how we implement the foundations of recursive analysis and topology based on open sets instead of points.

However, the introduction of nuclei with this definition *after* the lattice theory in the previous section is the reason why our discussion of open and closed subspaces and the Phoa principle in Remark 4.11 was deductively unsatisfactory. From a foundational point of view, the monadic structure should really come *first*. Then the Euclidean principle ($\sigma \wedge F\sigma \Leftrightarrow \sigma \wedge F\top$), which is the part of the Phoa principle that applies to intuitionistic set theory or the subobject classifier in an elementary topos, is exactly what is required to make $E\phi \equiv \phi \wedge \theta$ a nucleus in the abstract sense, or equivalently to justify the first of the two Gentzen-style rules. This is essentially how this principle was discovered [C, §3].

The results that we have sketched here are treated in greater detail in the earlier papers that developed ASD. The foundations of the theory, together with the methodology and motivations behind it, are surveyed in [O].

6 Dedekind cuts

We now return to Richard Dedekind's construction of the real line, which we introduced in Section 2. Whilst this is very familiar, we need to be sure that it can be carried out in our very weak logic. We begin with what we require of the rationals for the purposes of topology, postponing arithmetic to Sections 11–13. Both the ingredients (Q) and the result (R) of the construction satisfy the first definition, but we shall use combinatorial operations on the rationals which will rely on decidable equality (Remark 6.15).

Definition 6.1 A ***dense linear order without endpoints*** is an overt Hausdorff object Q (Definitions 4.12 and 4.14) with an open binary relation $<$ that is, for $p, q, r : Q$,

(a) transitive and interpolative (dense):	$(p < r)$	$\iff$	$(\exists q.\, p < q < r)$
(b) extrapolative (without endpoints):	$(\exists p.\, p < q)$	$\iff$	$\top \iff (\exists r.\, q < r)$
(c) linear (total or trichotomous):	$(p \neq q)$	$\iff$	$(p < q) \vee (q < p)$.

The $\exists q$ in the interpolation axiom does not necessarily mean the midpoint $\frac{1}{2}(p + r)$. It might instead mean the "simplest" interpolant, in some sense, as in John Conway's number system [Con76]. In practical computation, the best choice of q may be determined by considerations elsewhere in the problem, *cf.* unification in PROLOG (Remark 3.4).

Lemma 6.2

$$\begin{aligned} (p < q) \ \wedge \ (s < t) &\implies (p < t) \ \vee \ (s < q) \\ (p < q) \ \wedge \ (q < r) &\implies (p < r) \implies (p < q) \ \vee \ (q < r) \end{aligned}$$

□

Lemma 6.3 If Q is discrete then $=$ and $<$ are decidable, with $(p \not< q) \equiv (p = q) \vee (q < p)$. This is the case when Q is countable (by which we mean strictly that $Q \cong \mathbb{N}$, Remark 5.22), in which case the usual order on $\mathbb{N}$ induces another (well founded) order $\prec$ on Q that we call ***simplicity***. □

Examples 6.4 Such countable Q may consist of
(a) all fractions n/m with $m \neq 0$, where "simpler" fractions have smaller denominators;
(b) dyadic or decimal fractions $k/2^n$ and $k/10^n$;
(c) finite continued fraction expansions; or
(d) roots of polynomials with integer coefficients, where the notion of simplicity is given by the degree and coefficients of the polynomials;

where in all cases $<$ is the arithmetical order. For the sake of motivation and of having a name for them, we shall refer to the elements of *any* such countable Q as ***rationals***, but for a particular target computation any of the above structures may be chosen. We keep the dyadic rationals in mind for practical reasons, but they have no preferred role in the definition.

Proposition 6.5 Any two countable dense linear orders without endpoints are order-isomorphic.

Proof The isomorphism $Q_1 \cong Q_2$ is built up from finite lists of pairs (Theorem 5.21) by course-of-values recursion and definition by description (Axiom 4.24). Given such a list, we may add another pair that consists of the simplest absentee from Q_1, together with its simplest order-match from Q_2. At the next stage, we start with an absentee from Q_2. □

Remark 6.6 Of course, this well known model-theoretic result misses the point of real analysis. When we apply it to different notions of "rationals", it destroys the arithmetic operations and the Archimedean property (Axiom 11.2). In Definition 13.1 we shall show how order automorphisms (strictly monotone functions) of R arise from *relations* rather than functions on Q.

Remark 6.7 There are many formulations of Dedekind cuts. We choose a two-sided version that appears as Exercise 5.3.3 in [TvD88], but we don't know who originally formulated it. That book, like many other accounts, uses a one-sided version in its main development.

As in the case of open and closed subspaces (Remark 3.2f), the relationship between the two sides of the cut is *contravariant*. One-sided accounts obtain one from the other by means of (Heyting) negation — with some awkwardness, since arithmetic negation ought to be involutive, whereas logical negation is not, in intuitionistic logic.

The ASD calculus does not allow negation of predicates. Nor, indeed, is there any kind of contravariant operation, as it would be neither continuous nor computable. So we have to use both halves of the cut. But this is not just to overcome a technical handicap of our weak logic: like open and closed subspaces, the two parts play complementary roles that are plainly of equal importance. We also saw in Section 2 that it is natural to generalise Dedekind cuts by *not* forcing the halves to be complementary, the resulting space being the interval domain.

Definition 6.8 Formalising Remark 2.1, a (***Dedekind***) ***cut*** (δ, υ) in a (not necessarily discrete) dense linear order without endpoints $(Q, <)$ is a pair of predicates

$$\text{for}\quad d, u : Q, \qquad \delta d,\ \upsilon u : \Sigma,$$

possibly involving parameters, such that

$$\begin{array}{lcll}
\upsilon u & \Leftrightarrow & \exists t{:}Q.\ \upsilon t\ \wedge\ (t < u) & \upsilon \text{ rounded upper} \\
\delta d & \Leftrightarrow & \exists e{:}Q.\ (d < e)\ \wedge\ \delta e & \delta \text{ rounded lower} \\
\top & \Leftrightarrow & \exists u{:}Q.\ \upsilon u & \text{bounded above} \\
\top & \Leftrightarrow & \exists d{:}Q.\ \delta d & \text{bounded below} \\
\bot & \Leftrightarrow & \delta d\ \wedge\ (u < d)\ \wedge\ \upsilon u & \text{disjoint} \\
(d < u) & \Rightarrow & (\delta d\ \vee\ \upsilon u) & \text{order-located}
\end{array}$$

We call δ or υ ***rounded*** if they satisfy the first or second of the above properties, these being the most important. Their types are the ascending and descending reals ($\underline{R}$, $\overline{R}$, Example 2.8 and Definition 7.3). Lemma 7.5 explains why we have defined disjointness in this way, in preference to several others, and we shall discuss a stronger notion of locatedness that is needed for arithmetic in Proposition 11.15. When (δ, υ) are rounded, bounded and disjoint, they represent an interval (Definitions 2.3 and 7.8). We construct the space IR of intervals in the next section, *on the way* to the construction of R itself.

We do not introduce the "set" of cuts. However, we reformulate this definition using a little category theory to provide the context in which to understand both our construction in ASD and the alternatives in other foundational settings in Section 15.

Remark 6.9 The six axioms above are (essentially) *equations* (*cf.* Axiom 4.5). The boundedness axioms are already equations between terms of type Σ, whilst the roundedness and disjointness conditions become equations of type Σ^Q and $\Sigma^{Q\times Q}$ when we λ-abstract the variables. Finally, locatedness may be rewritten as

$$\big(\lambda du.\ \delta d \vee \upsilon u \vee (d < u)\big)\ =\ \big(\lambda du.\ \delta d \vee \upsilon u\big) \quad \text{of type} \quad \Sigma^{Q\times Q}.$$

We therefore expect R to be the subspace of $\Sigma^Q \times \Sigma^Q$ of those pairs (δ, υ) that satisfy these six equations. In the language of category theory, this is the equaliser (Definition 5.3) of the parallel pair of arrows that collects them together:

$$\begin{tikzcd}
Q \arrow[r, "j"] & R \arrow[r, tail, "i"] & \Sigma^Q \times \Sigma^Q \arrow[r, shift left, "\text{LHS}"] \arrow[r, shift right, "\text{RHS}"'] & \Sigma^Q \times \Sigma^Q \times \Sigma \times \Sigma \times \Sigma^{Q\times Q} \times \Sigma^{Q\times Q} \\
 & \Gamma \arrow[u, dotted, "a"] \arrow[ur, "(\delta,\upsilon)"'] & &
\end{tikzcd}$$

We shall explain how to construct this equaliser in ASD, *i.e.* to define the new type R of Dedekind cuts, in the next two sections. For the remainder of this one, Definition 6.8 is merely a list of properties that a pair of predicates on Q may or may not have.

Real numbers admit two order relations, which we shall now define in terms of cuts. Recall from Remark 4.6(b) that we expect $a < b$ to be a *predicate* (a term of type Σ) but $a \leq b$ to be a *statement* (an equation between such terms).

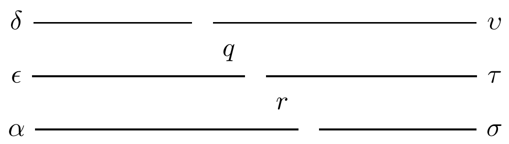
Definition 6.10 The irreflexive (strict) order relation on cuts is the predicate

$$(\delta, \upsilon) < (\epsilon, \tau) \;\equiv\; \exists q.\, \upsilon q \wedge \epsilon q.$$

The idea is that the right part of the smaller cut overlaps the left part of the larger.

Lemma 6.11 The relation $<$ is transitive.

Proof



$\square$

Lemma 6.12 As in Remark 2.1, each $q : Q$ gives rise to a cut $jq \equiv (\delta_q, \upsilon_q)$ with

$$\delta_q \;\equiv\; \lambda d.\,(d < q) \qquad \text{and} \qquad \upsilon_q \;\equiv\; \lambda u.\,(q < u).$$

Beware that q is a sub-script! This is an order embedding, in the sense that

$$(q < s) \;\Leftrightarrow\; (\exists r.\, q < r < s) \;\Leftrightarrow\; (\exists r.\, \upsilon_q r \wedge \delta_s r) \;\equiv\; \big((\delta_q, \upsilon_q) < (\delta_s, \upsilon_s)\big).$$

The map j makes Q an example of the test object Γ in the equaliser diagram above. Then the map $Q \to \Sigma^Q \times \Sigma^Q$ by $q \mapsto (\delta_q, \upsilon_q)$ is mono (1–1), as is $j : Q \to R$. $\square$

When we compare a real number (*i.e.* a cut) with a rational, we recover the lower or upper part of the cut:

Lemma 6.13 $\quad (\delta_d, \upsilon_d) < (\delta, \upsilon) \iff \delta d \quad$ and $\quad (\delta, \upsilon) < (\delta_u, \upsilon_u) \iff \upsilon u.$ $\square$

We would like to show next that $<$ is interpolative and extrapolative, but even to *say* that we need $\exists_R$, for which we have to define R itself as a type in ASD, so we leave this until Section 9.

Now we consider the *reflexive* (non-strict) order $\leq$. This is given by implication or inclusion between the cuts considered as predicates or subsets.

Proposition 6.14 Let (δ, υ) and (ϵ, τ) be cuts, possibly involving parameters. Then the three statements

$$\delta \geqslant \epsilon : \Sigma^Q, \qquad \upsilon \leqslant \tau : \Sigma^Q \qquad \text{and} \qquad \big((\delta, \upsilon) < (\epsilon, \tau)\big) \;\Leftrightarrow\; \bot$$

are equivalent. We write $(\epsilon, \tau) \leq (\delta, \upsilon)$ for any of them. Notice that the arithmetical order $\leq$ agrees with the logical one $\leqslant$ for the ascending reals δ and ϵ, but is the reverse for the descending reals τ and υ, *cf.* Example 2.8.

Proof Since $\big((\delta, \upsilon) < (\epsilon, \tau)\big)$ is $\exists q.\, \upsilon q \wedge \epsilon q$, it is $\bot$ if either $\epsilon \leqslant \delta$ or $\upsilon \leqslant \tau$, since cuts are rounded and disjoint.

Conversely, by the definition of $\exists$, another way of writing the statement with $<$ is

$$\text{for any} \quad q : Q, \qquad (\upsilon q \wedge \epsilon q) \;\Rightarrow\; \bot$$

Then, for $p : Q$,

$$\begin{array}{rcll} \epsilon p & \Leftrightarrow & \exists q.\,(p < q) \wedge \epsilon q & \epsilon \text{ is rounded} \\ & \Rightarrow & \exists q.\,(\delta p \vee \upsilon q) \wedge \epsilon q & (\delta, \upsilon) \text{ located} \\ & \Leftrightarrow & \exists q.\,(\delta p \wedge \epsilon q) \;\vee\; (\upsilon q \wedge \epsilon q) & \text{distributivity} \\ & \Rightarrow & \exists q.\,(\delta p \vee \bot) \;\equiv\; \delta p & \text{hypothesis,} \end{array}$$

and similarly $\upsilon p \Rightarrow \tau p$. $\square$

There is one more order-theoretic property that we shall need: it develops the notion of order-locatedness a little towards arithmetic, for which the stronger definition will be needed. The fact that this Lemma does not require disjointness, and so also applies to the back-to-front intervals in Remark 2.20, will become relevant when we study the existential quantifier.

Remark 6.15 We present this proof in the usual informal style, in which "$\cdots$" indicates the variable length of the list, but the formal proof, of course, uses list induction. In particular, the ordering of the list is an open predicate that, like the n-fold conjunction, is easily definable by recursion on lists. This means that we rely on the methods of "discrete mathematics" that were outlined in and following Theorem 5.21, but these only work for *overt discrete* spaces. On the other hand, for Definition 6.1 it had to be Hausdorff, so **equality on Q must be decidable**, *cf.* Lemmas 4.13 and 6.3.

Lemma 6.16 Let $q_0 < q_1 < \cdots < q_n$ be a strictly ascending finite sequence of rationals, and let (δ, υ) be rounded, located and bounded, with δq_0 and υq_n. Then this (pseudo-)cut belongs to at least one of the *overlapping* open intervals

$$(q_0, q_2), \quad (q_1, q_3), \quad \ldots, \quad (q_{n-2}, q_n),$$

in the sense that

$$\bigvee_{j=0}^{n-2} (\delta q_j \wedge \upsilon q_{j+2}).$$

Proof The hypotheses δq_0 and υq_n and order-locatedness of (δ, υ) applied to each pair $q_j < q_{j+1}$ give

$$\delta q_0 \ \wedge\ (\delta q_0 \vee \upsilon q_1) \ \wedge\ (\delta q_1 \vee \upsilon q_2) \ \wedge\ \cdots\ \wedge\ (\delta q_{n-1} \vee \upsilon q_n) \ \wedge\ \upsilon q_n,$$

which we expand using distributivity. Of the 2^n disjuncts, $n+1$ yield those in the claim, including two repetitions:

$$\begin{array}{lcl} \delta q_0 \wedge \upsilon q_1 \wedge \upsilon q_2 \wedge \cdots \wedge \upsilon q_n & \Rightarrow & \delta q_0 \wedge \upsilon q_2 \\ \delta q_0 \wedge \cdots \wedge \delta q_j \wedge \upsilon q_{j+2} \wedge \cdots \wedge \upsilon q_n & \Rightarrow & \delta q_j \wedge \upsilon q_{j+2} \\ \delta q_0 \wedge \cdots \wedge \delta q_{n-2} \wedge \delta q_{n-1} \wedge \upsilon q_n & \Rightarrow & \delta q_{n-2} \wedge \upsilon q_n. \end{array}$$

Each of the remaining $2^n - n - 1$ disjuncts consists of a conjunction of δs up to some δq_j and of υs starting from υq_i, with $1 \le i \le j+1 \le n$. Such a conjunction vanishes if δ and υ are disjoint. But even if they are not, since υ is upper we have

$$\delta q_j \wedge \upsilon q_i \implies \delta q_j \wedge \upsilon q_{j+2},$$

which is one of the disjuncts in the claim. $\square$

Exercise 6.17 Formulate the following in accordance with Remark 4.6, and prove them:
(a) if $a \le b < c \le d$ then $a < d$;
(b) if $a < b$ or $a = b$ then $a \le b$;
(c) if $a \le b$ and $a \ne b$ then $a < b$;
(d) if δd then $(d < u) \vee \delta u$;
(e) if $(d < u) \vee \delta u$ for all u then δd [NB: $\forall u : Q$ is not allowed];
(f) not all $d : Q$ satisfy δd.
Finally, verify that the results about $<$ and $\le$ make no use of the boundedness axiom for cuts, and therefore also apply to $+\infty \equiv (\top, \bot)$ and $-\infty \equiv (\bot, \top)$. $\square$

7 The interval domain in ASD

You might reasonably expect that our next task would be to construct R as an object of ASD. This would involve showing that the formula $\mathcal{E}$ in Section 2 satisfies the definition of a nucleus in Section 5, and that its admissible terms are exactly the Dedekind cuts of Section 6.

However, hard experience with nuclei has shown that one cannot go at them like a bull in a china shop: whilst the calculations are elementary, they are very tricky. In fact, it is precisely because we want to give an *elementary* proof (in the technical sense of the word, *i.e.* in a very weak logic) that it is *difficult*. If we were willing to assume all of the structure of set theory, Proposition 2.17 for $\mathbb{R}$ would already have justified the equations in Lemma 5.4. The elementary proof requires messy manipulation of sequences of rationals.

Sorting out the mess teaches us two things conceptually.

One is that we need to construct the ascending and descending reals and the interval domain first. The other is that the Euclidean nucleus $\mathcal{E}$ turns out to be a composite of four parts that are examples of the easy ways of defining nuclei in ASD, namely idempotents, open subspaces, closed subspaces and Scott-continuous localic nuclei (Definition 5.5 and Lemma 5.9). As we shall also see with arithmetic, the interval domain collects the "easy" properties of real numbers, clearing the ground for work on the difficult property of locatedness.

The technical difficulties that we resolve in this section may seem to be of our own making, as we have allowed $\delta, \upsilon : \Sigma^Q$ to be *arbitrary* λ-terms of this type. These terms may, for example, pick out individual rationals for "special" treatment, in the way that we find in the pathological counterexamples of real analysis but not its fundamental intuitions. As we explained in Remark 2.13, we start from Σ^Q because that is as close to computation as is reasonable, whereas the other structures are partial compromises with mathematics. (The order and arithmetic on Q are complicated when encoded *via* the bijection $Q \cong \mathbb{N}$, but are well known to lie within a weak fragment of logic like ASD.) Also, the types in ASD must be defined progressively from $\mathbf{1}$ and $\mathbb{N}$ using $\times$, $\Sigma^{(-)}$ and Σ-split subspaces, so we have to use $\Sigma^Q(\cong \Sigma^{\mathbb{N}})$ on the way to constructing the interval domain and finally R.

We dispose of most of the arbitrariness by constructing the spaces of ascending and descending reals (Example 2.8) as retracts of Σ^Q.

Notation 7.1 For arbitrary $\delta, \upsilon : \Sigma^Q$ and $\Phi : \Sigma^{\Sigma^Q \times \Sigma^Q}$, define

$$\hat{\delta} d \;\equiv\; \exists e.\, (d < e) \wedge \delta e, \qquad \check{\upsilon} u \;\equiv\; \exists t.\, \upsilon t \wedge (t < u),$$

$$B(\delta, \upsilon) \;\equiv\; \exists du.\, \delta d \wedge \upsilon u, \qquad \nabla(\delta, \upsilon) \;\equiv\; \exists du.\, \delta d \wedge (u < d) \wedge \upsilon u$$

and
$$\mathcal{E}_1 \Phi(\delta, \upsilon) \;\equiv\; \exists d < u.\, \hat{\delta} d \wedge \check{\upsilon} u \wedge \Phi(\delta_d, \upsilon_u).$$

As the name "roundedness" suggests, $\hat{\delta}$ and $\check{\upsilon}$ trim off the endpoints of cuts of the kind that Dedekind originally introduced, *cf.* Remark 2.1, as well as closing the subsets down- or upwards. The predicates B and ∇ capture boundedness and (non-)disjointness respectively (note the $u < d$ in ∇!). We shall find *three* nuclei that define the interval domain, but $\mathcal{E}_1$ will turn out to be the most useful one.

Lemma 7.2 These operations $\delta \mapsto \hat{\delta}$ and $\upsilon \mapsto \check{\upsilon}$ are *idempotent*. They fix δ and υ iff these are rounded (Definition 6.8). In particular, δ_d, $\top$ and $\bot$ are fixed by the first operation (rounding downwards), whilst υ_u, $\top$ and $\bot$ are fixed by the second (upwards). $\square$

Definition 7.3 Using Lemma 5.9, these idempotents on Σ^Q define types

$$\hat{-} : \underline{R} \lhd \Sigma^Q \quad \text{and} \quad \check{-} : \overline{R} \lhd \Sigma^Q,$$

called the ***ascending*** and ***descending reals*** respectively. The product $\underline{R} \times \overline{R}$ is similarly a retract of $\Sigma^Q \times \Sigma^Q$, so it is given by the nucleus

$$\mathcal{E}_{\mathsf{r}}\Phi(\delta, \upsilon) \equiv \Phi(\hat{\delta}, \check{\upsilon}) \equiv \Phi(\lambda d.\, \exists e.\, d < e \wedge \delta e,\ \lambda u.\, \exists t.\, \upsilon t \wedge t < u).$$

We may think of the lattice $\underline{R} \times \overline{R}$ as the space of "intervals" that may be infinite and even back-to-front (Remark 2.20).

Lemma 6.12 defines "inclusion" maps $\delta_{(-)} : Q \to \underline{R} \lhd \Sigma^Q$ and $\upsilon_{(-)} : Q \to \overline{R} \lhd \Sigma^Q$. These can be extended to $R \to \underline{R}$ and $R \to \overline{R}$, of course simply as the projections that extract δ and υ from the cut $(\delta, \upsilon) : R$. However, by monotonicity (Axiom 4.8), there is no map

$$\underline{R} \to \overline{R}, \qquad \overline{R} \to \underline{R}, \qquad \underline{R} \to R \qquad \text{or} \qquad \overline{R} \to R$$

that commutes with these inclusions. This is why both δ and υ are needed in the definition of Dedekind cuts.

Exercise 7.4 Construct the ascending natural numbers, with ∞ as top, in a similar way.

Next we consider disjointness. Of the three possible definitions, in Definition 6.8 we chose the one that is invariant under rounding, so that the nuclei commute, *cf.* Lemma 5.11.

Lemma 7.5 Let $\delta, \upsilon : \Sigma^Q$. Then $\nabla(\delta, \upsilon) \Leftrightarrow \nabla(\hat{\delta}, \check{\upsilon})$ and $\mathcal{E}_1\nabla(\delta, \upsilon) \Leftrightarrow \bot$.
Also, the statements (a,d,e,f) below are equivalent:

$$\begin{array}{llll@{\qquad}llll}
\text{(a)} & \nabla(\delta, \upsilon) & \Leftrightarrow & \bot & \text{(d)} & \nabla(\hat{\delta}, \check{\upsilon}) & \Leftrightarrow & \bot \\
\text{(b)} & \delta d \wedge \upsilon u & \Rightarrow & d < u & \text{(e)} & \hat{\delta} d \wedge \check{\upsilon} u & \Rightarrow & d < u \\
\text{(c)} & \delta q \wedge \upsilon q & \Leftrightarrow & \bot & \text{(f)} & \hat{\delta} q \wedge \check{\upsilon} q & \Leftrightarrow & \bot
\end{array}$$

However, if Q is discrete then $\delta \equiv \lambda d.\,(d \le r)$ and $\upsilon \equiv \lambda u.\,(r \le u)$ satisfy (a,d,e,f) but not (b,c); in fact (b) is equivalent to (a&c).

Proof

$$\begin{array}{rcl}
\nabla(\hat{\delta}, \check{\upsilon}) & \equiv & \exists detu.\, \delta e \wedge (t < u < d < e) \wedge \upsilon t \\
& \Leftrightarrow & \exists et.\, \delta e \wedge (t < e) \wedge \upsilon t \ \equiv\ \nabla(\delta, \upsilon) \\
\mathcal{E}_1\nabla(\delta, \upsilon) & \equiv & \exists d < u.\, \hat{\delta} d \wedge \check{\upsilon} u \wedge \nabla(\delta_d, \upsilon_u) \\
& \equiv & \exists d < u.\, \hat{\delta} d \wedge \check{\upsilon} u \wedge \exists et.\, \delta_d e \wedge (t < e) \wedge \upsilon_u t \\
& \Rightarrow & \exists detu.\, d < u < t < e < d \ \Rightarrow\ \bot.
\end{array}$$

□

Definition 7.6 The ***unbounded interval domain*** is the closed subspace $IR^\infty \subset \underline{R} \times \overline{R}$ that is co-classified by ∇, so it consists of those (δ, υ) for which $\nabla(\delta, \upsilon)$ is *false* (Remark 4.11), *i.e.* which are rounded and disjoint. Using Lemma 5.9, $IR^\infty \subset \Sigma^Q \times \Sigma^Q$ is defined by the nucleus

$$\mathcal{E}_{\mathsf{rd}}\Phi(\delta, \upsilon) \equiv \nabla(\hat{\delta}, \check{\upsilon}) \vee \Phi(\hat{\delta}, \check{\upsilon}).$$

The inclusion map $IR^\infty \to \underline{R} \times \overline{R}$ provides the endpoints of any interval, but as ascending and descending reals: there is no such map $IR^\infty \to R \times R$ (*cf.* Remark 2.9). On the other hand, there is of course a function from $\{(d, u) \mid d \le u\} \lhd R \times R$ to IR^∞ that provides the interval with *specified* endpoints.

That version of the interval domain included $(\bot, \upsilon_u) \equiv [-\infty, u]$, $(\delta_d, \bot) \equiv [d, +\infty]$ and $(\bot, \bot) \equiv [-\infty, +\infty]$. The following construction, on the other hand, eliminates them; since it therefore has no least element, it is called a *pre*domain. It is again invariant under rounding.

Lemma 7.7 $\quad \nabla(\delta, \upsilon) \Longrightarrow B(\delta, \upsilon) \iff B(\hat{\delta}, \check{\upsilon}) \iff \mathcal{E}_1\top(\delta, \upsilon).$

Proof

$$\begin{aligned} B(\hat{\delta}, \check{\upsilon}) &\equiv \exists du.\, \hat{\delta}d \wedge \check{\upsilon}u \Leftrightarrow \exists d < u'.\, \hat{\delta}d \wedge \check{\upsilon}u' \equiv \mathcal{E}_1\top(\delta, \upsilon) \qquad u' > d, u \\ &\equiv \exists detu.\, (d < e) \wedge \delta e \wedge \upsilon t \wedge (t < u) \\ &\Leftrightarrow \exists et.\, \delta e \wedge \upsilon t \equiv B(\delta, \upsilon). \end{aligned}$$

□

Definition 7.8 The ***bounded interval predomain*** is the *open* subspace $IR \subset IR^\infty$ classified by B, so it consists of those (δ, υ) for which $B(\delta, \upsilon)$ is *true*. By Lemma 5.9, $IR \subset \Sigma^Q \times \Sigma^Q$ is defined by the nucleus

$$\mathcal{E}_{\mathsf{rdb}}\Phi(\delta, \upsilon) \equiv \nabla(\hat{\delta}, \check{\upsilon}) \vee \Phi(\hat{\delta}, \check{\upsilon}) \wedge B(\hat{\delta}, \check{\upsilon}),$$

which may be bracketed either way, since $\nabla \leqslant B$.

Warning 7.9 Here we are selecting the bounded pseudo-cuts from spaces of not necessarily bounded ones. Nuclei that do this do not preserve $\top$. This is related, *via* Lemma 5.13, to the fact that $[d, u]$ is compact, but R is not.

The following rule of inference will simplify matters a lot in the next section. Abstractly, it is just (monotonicity of) the I operator in Section 5 for the subspace $IR \subset \Sigma^Q \times \Sigma^Q$.

Proposition 7.10 In proving any implication $\Phi \leqslant \Psi$, we may assume that the pseudo-cuts to which they are applied are rounded, bounded and disjoint (or even that they have rational endpoints), at the cost of applying $\mathcal{E}_1$ to the result, in the sense that

$$\frac{\Gamma,\ (\delta, \upsilon) : IR \ \vdash\ \Phi(\delta, \upsilon) \ \Rightarrow\ \Psi(\delta, \upsilon)}{\Gamma,\ \delta, \upsilon : \Sigma^Q \ \vdash\ \mathcal{E}_1\Phi(\delta, \upsilon) \ \Rightarrow\ \mathcal{E}_1\Psi(\delta, \upsilon).}$$

Proof If $d < u$ then (δ_d, υ_u) is rounded, bounded and disjoint, so

$$\begin{aligned} \mathcal{E}_1\Phi(\delta, \upsilon) &\equiv \exists d < u.\, \hat{\delta}d \wedge \check{\upsilon}u \wedge \Phi(\delta_d, \upsilon_u) \\ &\Rightarrow \exists d < u.\, \hat{\delta}d \wedge \check{\upsilon}u \wedge \Psi(\delta_d, \upsilon_u) \equiv \mathcal{E}_1\Psi(\delta, \upsilon). \end{aligned}$$

□

Exercise 7.11 Use the following result to sharpen this rule to show that the premise

$$\Gamma,\ q : Q \ \vdash\ \Phi(\delta_q, \upsilon_q) \ \Rightarrow\ \Psi(\delta_q, \upsilon_q)$$

is sufficient to deduce $\mathcal{E}_1\Phi \leqslant \mathcal{E}_1\Psi$, and indeed $\mathcal{E}\Phi \leqslant \mathcal{E}\Psi$, as we shall see in the next section. □

Next we invoke Scott continuity of Φ, *cf.* Lemma 2.16.

Proposition 7.12 If (δ, υ) are rounded and bounded then, for any $\Phi : \Sigma^{\Sigma^Q \times \Sigma^Q}$,

$$\Phi(\delta, \upsilon) \iff \exists du.\, \delta d \wedge \upsilon u \wedge \Phi(\delta_d, \upsilon_u).$$

If (δ, υ) are also disjoint then $\Phi(\delta, \upsilon) \iff \exists d < u.\, \hat{\delta}d \wedge \check{\upsilon}u \wedge \Phi(\delta_d, \upsilon_u) \equiv \mathcal{E}_1\Phi(\delta, \upsilon)$.

Proof We have to show that the joins in Notation 7.1,

$$\hat{\delta} = \exists d.\, \delta d \wedge (\lambda x.\, x < d) \quad \text{and} \quad \check{\upsilon} = \exists u.\, \upsilon u \wedge (\lambda x.\, u < x),$$

are *directed* in the sense of Definition 4.20, so consider

$$N \equiv Q, \quad \alpha_d \equiv \delta d \quad \text{and} \quad \phi^d \equiv \delta_d \equiv \lambda x.\, x < d.$$

Then $\exists d.\, \alpha_d \Leftrightarrow \top$ by boundedness, and the binary operation $@ \equiv \mathsf{max}$ satisfies

$$\alpha_{d@e} \;\equiv\; \delta\big(\mathsf{max}(d,e)\big) \iff \delta d \wedge \delta e \;\equiv\; \alpha_d \wedge \alpha_e$$

as δ is lower, and $\quad \phi^{d@e}x \;\equiv\; x < \mathsf{max}(d,e) \;\Longleftarrow\; (x<d) \vee (x<e) \;\equiv\; (\phi^d \vee \phi^e)x.$
For the second, we use $@ \equiv \mathsf{min}$ instead. Then we deduce the result for rounded and bounded $(\delta,\upsilon) = (\hat{\delta},\check{\upsilon})$ from Axiom 4.21. If they are also disjoint then $\delta d \wedge \upsilon u \Rightarrow d < u$ by Lemma 7.5. $\square$

It may perhaps be surprising that this is the (one) place where Scott continuity is used in the construction, given that the manipulation of finite lists will happen in the next section. The reason is that this is the section in which we do *domain theory*. The intervals with *rational* endpoints provide a ***basis*** for the domain of *general* intervals in the sense of Remark 2.9, and the Proposition shows how we extend the definition of Φ from the basis to the whole domain.

In fact, we have shown that there is *at most one* extension. We need another condition to ensure that the extension *exists*, *i.e.* that we recover the given formula when we restrict back to the basis. This requirement is more than monotonicity with respect to interval inclusions, since the bounded interval predomain is a continuous but not algebraic dcpo.

Proposition 7.13 Let $\phi(d,u)$ be any predicate on $d,u : Q$, which are intended to be the endpoints of an interval. Then

$$\Phi(\delta,\upsilon) \;\equiv\; \exists du.\, \delta d \wedge \upsilon u \wedge \phi(d,u)$$

extends $\phi : Q \times Q \to \Sigma$ to $\Phi : \Sigma^Q \times \Sigma^Q \to \Sigma$, in the sense that $\Phi(\delta_d, \upsilon_u) \Leftrightarrow \phi(d,u)$, iff

$$\phi \quad \text{is } \textbf{\textit{rounded}}, \quad \phi(e,t) \;\Leftrightarrow\; \exists du.\, (d<e) \wedge (t<u) \wedge \phi(d,u).$$

If ϕ satisfies the same condition in just the restricted case where $d \leq u$, then it extends to the bounded interval predomain, $\Phi : \mathit{IR} \to \Sigma$, *i.e.* to rounded, bounded, disjoint pseudo-cuts (δ,υ). $\square$

We shall use this condition in the definition of the nucleus $\mathcal{E}$ in the next section, and again for arithmetic in Sections 11–13. Here we may complete the construction of the (bounded) interval (pre)domain in ASD:

Theorem 7.14 $\mathcal{E}_1$ is a nucleus (Definition 5.5), and (δ,υ) is admissible for it (Definition 5.7) iff it is rounded, bounded and disjoint, so $\mathit{IR} \cong \{\Sigma^{\Sigma^Q \times \Sigma^Q} \mid \mathcal{E}_1\}$.

Proof If (δ,υ) are rounded, bounded and disjoint, $\Phi(\delta,\upsilon) \Leftrightarrow \mathcal{E}_1\Phi(\delta,\upsilon)$ by Proposition 7.12, *i.e.* (δ,υ) is *admissible*. On the same hypothesis, therefore,

$$(\Phi \wedge \Psi)(\delta,\upsilon) \iff (\mathcal{E}_1\Phi \wedge \mathcal{E}_1\Psi)(\delta,\upsilon).$$

We deduce from this, using Proposition 7.10, that, for arbitrary (δ,υ),

$$\mathcal{E}_1(\Phi \wedge \Psi)(\delta,\upsilon) \iff \mathcal{E}_1(\mathcal{E}_1\Phi \wedge \mathcal{E}_1\Psi)(\delta,\upsilon),$$

which was one of the equations for a *nucleus*. We can prove the $\vee$-equation in the same way, or by observing that $\mathcal{E}_1$ is idempotent and preserves $\vee$.

Conversely, we have to deduce roundedness, boundedness and disjointness from instances of admissibility with respect to the new formula for $\mathcal{E}_1$, using carefully chosen Φ, which are provided by Lemmas 7.5 and 7.7.

For *disjointness*, let $\Phi \equiv \nabla$, so by admissibility, $\nabla(\delta,\upsilon) \;\Leftrightarrow\; \mathcal{E}_1\nabla(\delta,\upsilon) \;\Leftrightarrow\; \bot$.

For *boundedness*, let $\Phi \equiv \lambda\alpha\beta.\,\top$, so $B(\delta,\upsilon) \;\Leftrightarrow\; \mathcal{E}_1\Phi(\delta,\upsilon) \;\Leftrightarrow\; \Phi(\delta,\upsilon) \;\Leftrightarrow\; \top$.

For *roundedness*, consider $\Phi \equiv \lambda\alpha\beta.\, \alpha c \wedge \beta w$ with $c, w : Q$, so

$$\begin{array}{lrcl} & \Phi(\delta_d, \upsilon_u) & \Leftrightarrow & \delta_d c \wedge \upsilon_u w \;\Leftrightarrow\; (c < d) \wedge (u < w). \\ \text{Then} & \mathcal{E}_1\Phi(\delta, \upsilon) & \Leftrightarrow & \exists d < u.\, \hat{\delta} d \wedge \check{\upsilon} u \wedge \Phi(\delta_d, \upsilon_u) \\ & & \Leftrightarrow & \exists du.\, \hat{\delta} d \wedge \check{\upsilon} u \wedge (c < d) \wedge (d < u) \wedge (u < w) \\ & & \Leftrightarrow & \hat{\delta} c \wedge \check{\upsilon} w \wedge (c < w) \;\Leftrightarrow\; \hat{\delta} c \wedge \check{\upsilon} w, \end{array}$$

since $(\hat{\delta}, \check{\upsilon})$ are rounded and disjoint. By admissibility,

$$\delta c \wedge \upsilon w \equiv \Phi(\delta, \upsilon) \Leftrightarrow \mathcal{E}_1\Phi(\delta, \upsilon) \Leftrightarrow \hat{\delta} c \wedge \check{\upsilon} w. \qquad \square$$

We conclude by looking at two other nuclei that also define IR.

Exercise 7.15 Show that

(a) $\mathcal{E}_{\mathsf{r}}$ (Definition 7.3), $\mathcal{E}_{\mathsf{d}}$ and $\mathcal{E}_{\mathsf{b}}$ are nuclei, where

$$\mathcal{E}_{\mathsf{r}}\Phi(\delta, \upsilon) \;\equiv\; \Phi(\hat{\delta}, \check{\upsilon}), \qquad \mathcal{E}_{\mathsf{d}}\Phi \equiv \Phi \vee \nabla \qquad \text{and} \qquad \mathcal{E}_{\mathsf{b}}\Phi \equiv \Phi \wedge B;$$

(b) (δ, υ) is admissible for $\mathcal{E}_{\mathsf{r}}$, $\mathcal{E}_{\mathsf{b}}$ or $\mathcal{E}_{\mathsf{d}}$ iff it is rounded, bounded or disjoint respectively (the last being in the sense that we used in Definition 6.8 and Lemma 7.5(a,d,e,f));

(c) $\mathcal{E}_{\mathsf{r}}$, $\mathcal{E}_{\mathsf{b}}$ and $\mathcal{E}_{\mathsf{d}}$ commute pairwise (hint: use Lemmas 7.5 and 7.7);

(d) the composite $\mathcal{E}_{\mathsf{rb}} \equiv \mathcal{E}_{\mathsf{r}} \cdot \mathcal{E}_{\mathsf{b}} \equiv \mathcal{E}_{\mathsf{b}} \cdot \mathcal{E}_{\mathsf{r}}$ is

$$\mathcal{E}_{\mathsf{rb}}\Phi(\delta, \upsilon) \iff \exists du.\, \hat{\delta} d \wedge \check{\upsilon} u \wedge \Phi(\delta_d, \upsilon_u) \iff \exists et.\, \delta e \wedge \upsilon t \wedge \Phi(\delta_e, \upsilon_t)$$

(hint: use Axiom 4.7 with $B(\delta, \upsilon) \Leftrightarrow \top$ and Proposition 7.12); and

(e) the composite $\mathcal{E}_{\mathsf{rdb}}$ (Definition 7.8) is $\mathcal{E}_{\mathsf{rdb}} = \nabla \vee \mathcal{E}_1$ (hint: use Axiom 4.7 with $\nabla(\delta, \upsilon) \Leftrightarrow \bot$ and Lemma 7.5).

We prefer $\mathcal{E}_1$ because it preserves $\bot$, whereas $\mathcal{E}_{\mathsf{rdb}}\bot = \nabla$. We use this property in Theorem 9.2 to construct the existential quantifier. $\square$

Exercise 7.16 Let $\mathcal{E}_1'\Phi(\delta, \upsilon) \equiv \exists d < u.\, \delta d \wedge \upsilon u \wedge \Phi(\delta_d, \upsilon_u)$ and show that

(a) $\mathcal{E}_1'$ is a nucleus;

(b) (δ, υ) is admissible for $\mathcal{E}_1'$ iff it is rounded, bounded and disjoint (hint: the only difficulty with copying the proof for $\mathcal{E}_1$ is that it only shows that admissible (δ, υ) are disjoint, bounded and satisfy $\delta \leqslant \hat{\delta}$ and $\upsilon \leqslant \check{\upsilon}$, so a little more work is needed to deduce roundedness);

(c) $\mathcal{E}_1'$ also satisfies $\mathcal{E}_1' \;=\; \mathcal{E}_1' \cdot \mathcal{E}_{\mathsf{r}} \;=\; \mathcal{E}_1' \cdot \mathcal{E}_{\mathsf{r}} \cdot \mathcal{E}_1' \;\leqslant\; \mathcal{E}_1 \;=\; \mathcal{E}_r \cdot \mathcal{E}_1' \;=\; \mathcal{E}_r \cdot \mathcal{E}_1' \cdot \mathcal{E}_{\mathsf{r}}$; but

(d) there are $\delta, \upsilon : \Sigma^Q$ for which $\mathcal{E}_1\top(\delta, \upsilon) \Leftrightarrow \top$ and $\mathcal{E}_1'\top(\delta, \upsilon) \Leftrightarrow \bot$ (hint: *cf.* Lemma 7.7).

We prefer $\mathcal{E}_1$ because it commutes with $\mathcal{E}_{\mathsf{r}}$ (*cf.* Lemma 5.11), so the Σ-splitting for $IR \rightarrowtail \Sigma^Q \times \Sigma^Q$ given by $\mathcal{E}_1$ restricts to $IR \rightarrowtail \underline{R} \times \overline{R}$. $\square$

8 The real line as a space in ASD

We are at last ready to construct R as a Σ-split subspace of the bounded interval predomain IR, and hence of $\Sigma^Q \times \Sigma^Q$.

Remark 8.1 Because of Lemma 5.11, the process of defining R by means of a nucleus $\mathcal{E}$ on IR, where IR is in turn defined by a nucleus $\mathcal{E}_1$ on $\Sigma^Q \times \Sigma^Q$, is the same as giving a nucleus $\mathcal{E}$ on the larger space for which

$$\mathcal{E} \cdot \mathcal{E}_1 \;=\; \mathcal{E} \;=\; \mathcal{E} \cdot \mathcal{E} \;=\; \mathcal{E}_1 \cdot \mathcal{E}.$$

Naturally, it will simplify the proof if we can restrict attention to $(\delta, \upsilon) : IR$, *i.e.* pseudo-cuts that are rounded, bounded and disjoint, thereby isolating locatedness.

However, another reason for considering R as a subspace of IR is that $\mathcal{E}$ then becomes a nucleus in the sense of both ASD and locale theory (Definition 5.5). The latter says that $\mathsf{id} \leqslant \mathcal{E} = \mathcal{E}^2$ and it preserves meets, so we have to prove that

$$\Phi(\delta, \upsilon) \;\Leftrightarrow\; \mathcal{E}_1\Phi(\delta, \upsilon) \;\Rightarrow\; \mathcal{E}\Phi(\delta, \upsilon) \quad \text{and} \quad \mathcal{E}\Phi(\delta, \upsilon) \;\wedge\; \mathcal{E}\Psi(\delta, \upsilon) \;\Rightarrow\; \mathcal{E}(\Phi \wedge \Psi)(\delta, \upsilon),$$

on the assumption that $(\delta, \upsilon) : IR$. (Only the last of these actually remains to be done.) We use Proposition 7.10 to eliminate this hypothesis, by applying $\mathcal{E}_1$. In Proposition 8.9 we shall deduce the equations in Lemma 5.4 from these results.

Notation 8.2 As the formula in Proposition 2.17 only involves variables ranging over Q, Σ^Q and $\Sigma^{\Sigma^Q \times \Sigma^Q}$, we may import it into the logic of ASD. However, we prefer to use

$$\mathcal{E}\Phi(\delta, \upsilon) \;\equiv\; \exists n \geq 1.\, \exists q_0 < \cdots < q_n.\;\; \hat{\delta} q_0 \wedge \check{\upsilon} q_n \;\wedge\; \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k},\; \upsilon_{q_{k+1}}),$$

with $\hat{\delta}$ and $\check{\upsilon}$ instead of δ and υ, and no $n \equiv 0$ term (which would be ∇). The reason is the same as for the choice of $\mathcal{E}_1$ in preference to $\mathcal{E}_1'$ or $\mathcal{E}_{\mathsf{rdb}}$, namely that it commutes with $\mathcal{E}_{\mathsf{r}}$ and preserves $\perp$ (Exercises 7.15f). Propositions 2.15 and 2.17 could be adapted to this choice by replacing I there with

$$IV \;\equiv\; \{(D, U) \mid \exists deut.\, d < e \wedge e \in D \wedge [d, u] \subset V \wedge t \in U \wedge t < u\}.$$

We have already said in Remark 6.15 that equality on Q must be decidable in order to use ASD's methods of "discrete mathematics". In particular, note that "$\exists q_0 < \cdots < q_n$" is existential quantification over $\mathsf{List}(Q)$, not just $n+1$-fold quantification over Q.

Notice that $\mathcal{E}_1$ is the $n \equiv 1$ disjunct of $\mathcal{E}$, so $\mathcal{E}_1 \leqslant \mathcal{E}$.

We begin with some lemmas for managing the sequence.

Lemma 8.3 For any (δ, υ), suppose $\mathcal{E}\Phi(\delta, \upsilon)$ holds by virtue of the sequence $q_0 < \cdots < q_n$, and that $q_0 \leq r \leq q_n$. Then $\mathcal{E}\Phi(\delta, \upsilon)$ also holds by virtue of the sequence obtained by inserting r.

Proof If $q_k < r < q_{k+1}$ then $\Phi(\delta_{q_k}, \upsilon_{q_{k+1}}) \;\Rightarrow\; \Phi(\delta_{q_k}, \upsilon_r) \wedge \Phi(\delta_r, \upsilon_{q_{k+1}})$. ☐

Lemma 8.4 For any Φ, and understanding that $q_0 \equiv d$ and $q_n \equiv u$, the predicate

$$\theta(d, u) \equiv \exists d < q_1 < \cdots < q_{n-1} < u.\; \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k}, \upsilon_{q_{k+1}})$$

is rounded, in the sense that $\theta(e, t) \iff \exists du.\, (d < e) \wedge (t < u) \wedge \theta(d, u) \Longrightarrow e < t$.

Proof First, $\theta(d, u) \Rightarrow d < u$ because there are $d < q_1 < \cdots < q_{n-1} < u$.

[$\Leftarrow$] If $\theta(d, u)$ holds by virtue of some sequence $d = q_0 < \cdots < q_n = u$, and $d < e < t < u$, then the same sequence with e and t in place of q_0 and q_n, but omitting any $q_i < e$ or $> u$, justifies $\theta(e, t)$.

[$\Rightarrow$] By Lemma 8.3, we may assume that $n \geq 3$ in the expansion of θ. The idea is to separate the left- and rightmost conjuncts from the rest, and then enlarge them using Proposition 7.12:

$$\begin{aligned}
\theta(d, u) \;&\Leftrightarrow\; \exists d < e < t < u.\, \Phi(\delta_d, \upsilon_e) \wedge \theta(e, t) \wedge \Phi(\delta_t, \upsilon_u) \\
&\Leftrightarrow\; \exists c < d < e < t < u < w.\, \Phi(\delta_c, \upsilon_e) \wedge \theta(e, t) \wedge \Phi(\delta_t, \upsilon_w) \\
&\Rightarrow\; \exists c < d < u < w.\, \theta(c, w).
\end{aligned}$$

☐

This was the condition in Proposition 7.13 for θ to extend from intervals with rational endpoints to more general rounded, bounded and disjoint ones:

Lemma 8.5 If $d < u$ then

$$\mathcal{E}\Phi(\delta_d, \upsilon_u) \iff \exists d < q_1 \cdots < q_{n-1} < u.\ \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k}, \upsilon_{q_{k+1}}).$$

Although $\theta(d, u) \Leftrightarrow \bot$ when $d \geq u$, we shall see in Propositions 8.12 and 10.5 that, in this case,

$$\mathcal{E}\Phi(\delta_x, \upsilon_x) \iff \Phi(\delta_x, \upsilon_x) \qquad \text{and} \qquad \mathcal{E}\Phi(\delta_d, \upsilon_u) \iff \exists x{:}[u, d].\ \Phi(\delta_x, \upsilon_x).$$

Proof The expansion of $\mathcal{E}\Phi(\delta_d, \upsilon_u)$ involves $\delta_d q_0 \wedge \upsilon_u q_n \equiv (q_0 < d < u < q_n)$. So, by Lemma 8.3, we may assume that d and u occur in the given sequence, say as $d \equiv q_i$, $u \equiv q_j$. But then $q_0, \ldots, q_{i-2}$ and $q_{j+2}, \ldots, q_n$ are redundant, so without loss of generality $d \equiv q_1$ and $u \equiv q_{n-1}$. Hence by Lemma 8.4,

$$\mathcal{E}\Phi(\delta_d, \upsilon_u) \iff \exists q_0 < d < u < q_n.\ \theta(q_0, q_n) \iff \theta(d, u). \qquad \square$$

We can now carry out the plan in Remark 8.1. The first two results say that the subspace R that will be defined by $\mathcal{E}$ is contained in that defined by $\mathcal{E}_1$, which is *IR*. Notice that the expansion of $\mathcal{E}_1 \cdot \mathcal{E}$ in the second result inherits the $\hat{\delta}$ and $\check{\upsilon}$ from $\mathcal{E}_1$, so they must also be used in the definition of $\mathcal{E}$ in order to obtain the equation $\mathcal{E}_1 \cdot \mathcal{E} = \mathcal{E}$.

Lemma 8.6 $\mathcal{E}(\mathcal{E}_1\Phi)(\delta, \upsilon) \iff \mathcal{E}\Phi(\delta, \upsilon)$ and $\mathcal{E}_1(\mathcal{E}\Phi)(\delta, \upsilon) \iff \mathcal{E}\Phi(\delta, \upsilon)$.

Proof The first part involves the conjuncts $\mathcal{E}_1\Phi(\delta_{q_k}, \upsilon_{q_{k+1}})$ that occur in the expansion of the formula for $\mathcal{E}(\mathcal{E}_1\Phi)(\delta, \upsilon)$. Such $(\delta_{q_k}, \upsilon_{q_{k+1}})$ are rounded, bounded and disjoint, so

$$\mathcal{E}_1\Phi(\delta_{q_k}, \upsilon_{q_{k+1}}) \iff \Phi(\delta_{q_k}, \upsilon_{q_{k+1}})$$

by Proposition 7.12. These are the same conjuncts that are used in the expansion of $\mathcal{E}\Phi(\delta, \upsilon)$.

In the second part, using Lemma 8.5,

$$\begin{aligned}
\mathcal{E}_1(\mathcal{E}\Phi)(\delta, \upsilon) &\equiv \exists d < u.\ \hat{\delta} d \wedge \check{\upsilon} u \wedge \mathcal{E}\Phi(\delta_d, \upsilon_u) \\
&\equiv \exists d < u.\ \hat{\delta} d \wedge \check{\upsilon} u \wedge \exists d < q_1 \cdots < q_{n-1} < u.\ \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k}, \upsilon_{q_{k+1}}) \\
&\equiv \mathcal{E}\Phi(\delta, \upsilon).
\end{aligned} \qquad \square$$

Lemma 8.7 $\mathcal{E}^2\Phi(\delta, \upsilon) \Longrightarrow \mathcal{E}\Phi(\delta, \upsilon)$.

Proof The outer $\mathcal{E}$ of $\mathcal{E}^2$ involves a sequence $q_0 < \cdots < q_n$ and conjuncts $\mathcal{E}\Phi(\delta_{q_k}, \upsilon_{q_{k+1}})$ for each $0 \leq k \leq n-1$. By Lemma 8.5, the expansion of each of these involves a sequence

$$q_k = r_{k,0} < \cdots < r_{k,m_k} = q_{k+1}$$

with conjuncts $\Phi(\delta_{r_{k,j}}, \upsilon_{r_{k,j+1}})$. All of these sequences can be concatenated, to yield one that justifies $\mathcal{E}\Phi(\delta, \upsilon)$. $\square$

The final preparatory result is the one that says that $\mathcal{E}$ is a nucleus on *IR* in the sense of locale theory, as well as that of ASD.

Proposition 8.8 If (δ, υ) are rounded, bounded and disjoint then

$$\mathcal{E}\top(\delta, \upsilon) \iff \top \qquad \text{and} \qquad \mathcal{E}\Phi(\delta, \upsilon) \wedge \mathcal{E}\Psi(\delta, \upsilon) \Longrightarrow \mathcal{E}(\Phi \wedge \Psi)(\delta, \upsilon).$$

Proof As (δ, υ) are rounded and bounded, $\top \Leftrightarrow \mathcal{E}_1 \top(\delta, \upsilon) \Rightarrow \mathcal{E}\top(\delta, \upsilon)$ by Lemma 7.7.

Suppose that $\mathcal{E}\Phi$ and $\mathcal{E}\Psi$ are justified by sequences $q_0 < \cdots < q_n$ and $r_0 < \cdots < r_m$ respectively. Let

$$d \equiv \mathsf{max}(q_0, r_0) \quad \text{and} \quad u \equiv \mathsf{min}(q_n, r_m),$$

so d is either q_0 or r_0, which both satisfy δ, so δd holds, as similarly does υu. Since (δ, υ) are rounded and disjoint, $d < u$ by Lemma 7.5. Now let

$$d \equiv s_0 < \cdots < s_\ell \equiv u$$

be the union of the given sequences $q_0 < \cdots < q_n$ and $r_0 < \cdots < r_m$, removing duplicates and those terms $< d$ or $> u$. Such outlying points were already redundant in the expansion of $\mathcal{E}\Phi(\delta, \upsilon)$ and of $\mathcal{E}\Psi(\delta, \upsilon)$, whilst by Lemma 8.3 we may adjoin the members of one sequence to the other. Hence the new combined sequence serves for both $\mathcal{E}\Phi(\delta, \upsilon)$ and $\mathcal{E}\Psi(\delta, \upsilon)$, and therefore for $\mathcal{E}(\Phi \wedge \Psi)(\delta, \upsilon)$ too. □

Proposition 8.9 $\mathcal{E}$ is a nucleus (Definition 5.5), as it satisfies, for $\Phi, \Psi : \Sigma^{\Sigma^Q \times \Sigma^Q}$,

$$\mathcal{E}(\Phi \wedge \Psi) = \mathcal{E}(\mathcal{E}\Phi \wedge \mathcal{E}\Psi) \quad \text{and} \quad \mathcal{E}(\Phi \vee \Psi) = \mathcal{E}(\mathcal{E}\Phi \vee \mathcal{E}\Psi).$$

Proof For $(\delta, \upsilon) : \mathit{IR}$, Proposition 7.12 gives $\Phi(\delta, \upsilon) \Leftrightarrow \mathcal{E}_1\Phi(\delta, \upsilon) \Rightarrow \mathcal{E}\Phi(\delta, \upsilon)$, so

$$(\Phi \wedge \Psi)(\delta, \upsilon) \Rightarrow (\mathcal{E}\Phi \wedge \mathcal{E}\Psi)(\delta, \upsilon) \qquad (\Phi \vee \Psi)(\delta, \upsilon) \Rightarrow (\mathcal{E}\Phi \vee \mathcal{E}\Psi)(\delta, \upsilon)$$

and $$\mathcal{E}\Phi(\delta, \upsilon) \vee \mathcal{E}\Psi(\delta, \upsilon) \Rightarrow \mathcal{E}(\Phi \vee \Psi)(\delta, \upsilon),$$

whilst $$\mathcal{E}\Phi(\delta, \upsilon) \wedge \mathcal{E}\Psi(\delta, \upsilon) \Rightarrow \mathcal{E}(\Phi \wedge \Psi)(\delta, \upsilon)$$

by Proposition 8.8. Then we use Proposition 7.10 to eliminate the hypothesis $(\delta, \upsilon) : \mathit{IR}$, and obtain the first inequality in each row:

$$\begin{array}{rcccccc}
\mathcal{E}_1(\Phi \wedge \Psi) & \leqslant & \mathcal{E}_1(\mathcal{E}\Phi \wedge \mathcal{E}\Psi) & \leqslant & \mathcal{E}(\mathcal{E}\Phi \wedge \mathcal{E}\Psi) & & \\
\mathcal{E}_1(\Phi \vee \Psi) & \leqslant & \mathcal{E}_1(\mathcal{E}\Phi \vee \mathcal{E}\Psi) & \leqslant & \mathcal{E}(\mathcal{E}\Phi \vee \mathcal{E}\Psi) & & \\
\mathcal{E}_1(\mathcal{E}\Phi \wedge \mathcal{E}\Psi) & \leqslant & \mathcal{E}_1(\mathcal{E}(\Phi \wedge \Psi)) & \leqslant & \mathcal{E}^2(\Phi \wedge \Psi) & \leqslant & \mathcal{E}(\Phi \wedge \Psi) \\
\mathcal{E}_1(\mathcal{E}\Phi \vee \mathcal{E}\Psi) & \leqslant & \mathcal{E}_1(\mathcal{E}(\Phi \vee \Psi)) & \leqslant & \mathcal{E}^2(\Phi \vee \Psi) & \leqslant & \mathcal{E}(\Phi \vee \Psi).
\end{array}$$

The second inequality follows from $\mathcal{E}_1 \leqslant \mathcal{E}$, and the third, $\mathcal{E}^2 \leqslant \mathcal{E}$, is Lemma 8.7. Finally, from each $\mathcal{E}_1\Theta \leqslant \mathcal{E}\Omega$, we deduce $\mathcal{E}\Theta = \mathcal{E}(\mathcal{E}_1\Theta) \leqslant \mathcal{E}^2\Omega \leqslant \mathcal{E}\Omega$ by Lemmas 8.6 and 8.7. □

This fulfils the obligation that we have before invoking the ASD technology in Section 5, so we can now introduce the type $R \equiv \{\Sigma^{\Sigma^Q \times \Sigma^Q} \mid \mathcal{E}\}$. We still have to characterise its admissible terms, *i.e.* to show that the equalisers in Remarks 5.2 and 6.9 are isomorphic.

The following is the main lemma to do this. In it, (δ, υ) need not be disjoint, so it may represent a back-to-front interval like those in Remark 2.20, which we shall use to compute existential quantifiers. In the proof, we have to switch back from using abutting closed intervals to overlapping open ones, *cf.* the proof of Proposition 2.17.

Lemma 8.10 If (δ, υ) is located then $\mathcal{E}\Phi(\delta, \upsilon) \iff \mathcal{E}_1\Phi(\delta, \upsilon)$.

Proof [$\Leftarrow$] is trivial. For [$\Rightarrow$], we use Proposition 7.12 to widen the argument of each conjunct Φ a little towards the right, from $[q_k, q_{k+1}]$ to $[q_k, r_{k+1}]$, for any $q_{k+1} < r_{k+1} < q_{k+2}$:

$$\begin{aligned} \mathcal{E}\Phi(\delta, \upsilon) &\equiv \exists q_0 < \cdots < q_n.\, \hat{\delta} q_0 \wedge \check{\upsilon} q_n \wedge \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k}, \upsilon_{q_{k+1}}) \\ &\Rightarrow \exists q_0 < r_0 < q_1 < r_1 < \cdots < q_n < r_n.\, \hat{\delta} q_0 \wedge \check{\upsilon} r_n \wedge \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k}, \upsilon_{r_{k+1}}), \end{aligned}$$

where $\check{\upsilon} q_n \Rightarrow \check{\upsilon} r_n$ because $\check{\upsilon}$ is upper. As $(\hat{\delta}, \check{\upsilon})$ is located with $\hat{\delta} q_0$ and $\check{\upsilon} r_n$, we may apply Lemma 6.16 to this sequence. Because of the alternation of letters, it gives

$$\bigvee_{k=0}^{n-1} (\hat{\delta} q_k \wedge \check{\upsilon} q_{k+1}) \;\vee\; \bigvee_{k=0}^{n-1} (\hat{\delta} r_k \wedge \check{\upsilon} r_{k+1}).$$

We already have the corresponding $\Phi(\delta_{q_k}, \upsilon_{q_{k+1}})$ from the original abutting expansion, whilst the one with overlapping intervals provides

$$\Phi(\delta_{q_k}, \upsilon_{r_{k+1}}) \implies \Phi(\delta_{r_k}, \upsilon_{r_{k+1}}),$$

since $\delta_{q_k} \leqslant \delta_{r_k}$. Hence, with $e \equiv q_k < t \equiv q_{k+1}$ or $e \equiv r_k < t \equiv r_{k+1}$, we have

$$\mathcal{E}\Phi(\delta, \upsilon) \implies \exists e < t.\, \hat{\delta} e \wedge \check{\upsilon} t \wedge \Phi(\delta_e, \upsilon_t) \;\equiv\; \mathcal{E}_1\Phi(\delta, \upsilon). \qquad \square$$

Lemma 8.11 If (δ, υ) is rounded and bounded, and $\mathcal{E}\Phi(\delta, \upsilon) \implies \mathcal{E}_1\Phi(\delta, \upsilon)$ for all Φ, then (δ, υ) is located.

Proof For any $d < u$, there is a sequence of rationals

$$q_0 < d < q_1 < u < q_2 \quad \text{with} \quad \delta q_0 \quad \text{and} \quad \upsilon q_2.$$

These numbers satisfy $\delta_{q_0} d \Leftrightarrow \bot$, $\upsilon_{q_1} u \Leftrightarrow \top$, $\delta_{q_1} d \Leftrightarrow \top$ and $\upsilon_{q_2} u \Leftrightarrow \bot$, so

$$\exists q_0 < q_1 < q_2.\;\; \delta q_0 \;\wedge\; \upsilon q_2 \;\wedge\; (\delta_{q_0} d \vee \upsilon_{q_1} u) \;\wedge\; (\delta_{q_1} d \vee \upsilon_{q_2} u).$$

This is the $n \equiv 2$ disjunct of $\mathcal{E}\Phi(\delta, \upsilon)$, where $\Phi(\alpha, \beta) \equiv \alpha d \vee \beta u$. By hypothesis, this implies $\mathcal{E}_1\Phi(\delta, \upsilon)$, where

$$\begin{aligned} \mathcal{E}_1\Phi(\delta, \upsilon) &\equiv \exists e < t.\, \delta e \wedge \upsilon t \wedge \Phi(\delta_e, \upsilon_t) \\ &\equiv \exists e < t.\, \delta e \wedge \upsilon t \wedge (d < e \vee t < u) \;\Rightarrow\; \delta d \vee \upsilon u. \end{aligned} \qquad \square$$

Proposition 8.12 (δ, υ) is a cut iff it is admissible for $\mathcal{E}$.

Proof If it's a cut then in particular it's rounded, bounded and disjoint, so $\Phi(\delta, \upsilon) \Leftrightarrow \mathcal{E}_1\Phi(\delta, \upsilon)$ by Proposition 7.12. As it's also located, $\mathcal{E}\Phi(\delta, \upsilon) \Leftrightarrow \mathcal{E}_1\Phi(\delta, \upsilon)$ by Lemma 8.10, so (δ, υ) is admissible.

Conversely, admissibility of (δ, υ) for $\mathcal{E}$ with respect to both Φ and $\mathcal{E}_1\Phi$, give

$$\Phi(\delta, \upsilon) \iff \mathcal{E}\Phi(\delta, \upsilon) \iff \mathcal{E}(\mathcal{E}_1\Phi)(\delta, \upsilon) \iff \mathcal{E}_1\Phi(\delta, \upsilon),$$

where the middle implication is Lemma 8.6. Hence (δ, υ) is also admissible for $\mathcal{E}_1$, so it is rounded, bounded and disjoint by Theorem 7.14. It is also located, by Lemma 8.11. $\square$

Theorem 8.13 The type $R \equiv \{\Sigma^{\Sigma^Q \times \Sigma^Q} \mid \mathcal{E}\}$ is the equaliser in Remark 6.9, *i.e.* its terms are Dedekind cuts, and the inclusion $R \rightarrowtail \Sigma^{\Sigma^Q \times \Sigma^Q}$ is Σ-split. $\square$

In the following two sections, we shall show that this construction does make the real line Dedekind-complete and the closed interval compact.

9 Dedekind completeness

In this section we shall prove Dedekind's main result, that any *cut* of R corresponds to an *element* of R, and not one of a yet more complicated structure. However, in order to say what a "Dedekind cut of R" is, we must finish the proof that its order (Definition 6.10) is dense without endpoints, *i.e.* that it satisfies Definition 6.1. This in turn requires that the object R be overt and Hausdorff.

In this section we shall be particularly careful to state the inclusion $j : Q \to R$ that we defined in Lemma 6.12 by $q \mapsto (\delta_q, \upsilon_q)$, although we usually elide it elsewhere.

Lemma 9.1 Any property Φ that is true of some real number (*i.e.* a cut $a \equiv (\delta, \upsilon)$), is also true of some interval $[e, t]$ around it, in the sense that

$$\begin{aligned} \Phi(\delta, \upsilon) \quad &\Leftrightarrow \quad \exists e < t{:}Q.\ \delta e \wedge \upsilon t \wedge \Phi(\delta_e, \upsilon_t) \\ &\Rightarrow \quad \exists e < t{:}Q.\ \delta e \wedge \upsilon t \wedge \Phi(\delta_e, \upsilon_e) \wedge \Phi(\delta_t, \upsilon_t). \end{aligned}$$

So in particular it is true of any sufficiently close rational e or t, *i.e.*

$$\Phi(ix) \ \equiv\ \Phi(\delta, \upsilon) \implies \exists q{:}Q.\ \Phi(\delta_q, \upsilon_q) \ \equiv\ \exists q{:}Q.\ \Phi(ijq).$$

Note, however, that we cannot *yet* write $\forall x{:}[e, t].\ \Phi(ix)$, but see Corollary 10.8.

Proof The first part is Proposition 7.12 in the case where (δ, υ) is a cut (rounded, bounded, disjoint and located). For the second part, $\delta_e \leqslant \delta_t$ and $\upsilon_t \leqslant \upsilon_e$. $\square$

Theorem 9.2 R is ***overt***, in which the ***existential quantifier*** satisfies the rule

$$\frac{\Gamma,\ x : R\ \vdash\ \phi x\ \Rightarrow\ \sigma}{\Gamma\ \vdash\ \exists x{:}R.\ \phi x\ \Rightarrow\ \sigma}$$

for any $\Gamma \vdash \phi : \Sigma^X,\ \sigma : \Sigma$, where

$$\exists x{:}R.\ \Phi(ix)\ \Leftrightarrow\ \exists q{:}Q.\ \Phi(ijq)\ \Leftrightarrow\ \mathcal{E}_1\Phi(\top, \top)\ \Leftrightarrow\ \mathcal{E}\Phi(\top, \top)\ \Rightarrow\ \Phi(\top, \top).$$

However, the converse of the last implication fails for $\Phi \equiv \nabla$ (Lemma 7.5).

Proof This is an application of Lemma 5.13, using the fact that $\Sigma^Q \times \Sigma^Q$ is a lattice and IR is the subspace of it that is defined by the nucleus $\mathcal{E}_1$, which preserves $\bot$.

Let's also give the proof directly. As Q is overt, we already have the rule with Q in place of R. The top line (for R) entails the similar judgement for Q by restriction, and then this entails the common bottom line. Conversely, by Lemma 9.1, $\phi x\ \Rightarrow\ \exists q{:}Q.\ \phi(jq)\ \Rightarrow\ \sigma$. Using it again,

$$\exists q{:}Q.\ \Phi(\delta_q, \upsilon_q)\ \Leftrightarrow\ \exists d < q < u.\ \Phi(\delta_d, \upsilon_u)\ \Leftrightarrow\ \exists d < u.\ \Phi(\delta_d, \upsilon_u)\ \equiv\ \mathcal{E}_1\Phi(\top, \top)$$

by Proposition 7.12.

Finally, we can use Lemma 8.10 to interchange $\mathcal{E}$ and $\mathcal{E}_1$ in the definition of $\exists$, that is, $\mathcal{E}\Phi(\top, \top) \Leftrightarrow \mathcal{E}_1\Phi(\top, \top)$, because $(\top, \top)$ is located. $\square$

Next we consider the application to R of the axioms that we introduced for Q in Section 6, namely dense linear orders, roundedness, boundedness, disjointness and locatedness. Their definitions involve existential quantification over R, but by Theorem 9.2 we may replace this by quantification over Q.

Theorem 9.3 R is ***Hausdorff*** (Definition 4.12), and $<$ (Definition 6.10) is a ***dense linear order without endpoints*** (Definition 6.1).

Proof By Proposition 6.14, if $a \equiv (\delta, \upsilon)$ and $b \equiv (\epsilon, \tau)$ are cuts then

$$\frac{a \;=\; b : R}{\big((a < b) \;\vee\; (b < a)\big) \;\Leftrightarrow\; \bot}$$

so R is Hausdorff, and the order satisfies the trichotomy law (Definition 6.1(c)); this is transitive by Lemma 6.11. It is extrapolative and interpolative in the sense that

$$\top \iff \exists du{:}Q.\ \delta d \wedge \upsilon u \iff \exists du.\,(\delta_d, \upsilon_d) < (\delta, \upsilon) < (\delta_u, \upsilon_u)$$

and
$$(\delta, \upsilon) < (\epsilon, \tau) \quad\Longrightarrow\quad \exists q{:}Q.\ (\delta, \upsilon) < (\delta_q, \upsilon_q) < (\epsilon, \tau),$$

but this existential quantification over Q may be considered as one over R. □

We can now define Dedekind cuts over R, *i.e.* to formulate Definition 6.8 and essentially to repeat the construction with R in place of Q. In doing so, we find once again that the ascending and descending reals are natural intermediate steps. As before, it is easy to define them, and to show that they are the same, whether we start from Q or from R, *cf.* Exercise 2.2. After that, we simply have to show that the other conditions carve out $R \subset \underline{R} \times \overline{R}$ in the same way in both versions.

Proposition 9.4 The maps $\Sigma^Q \rightleftarrows \Sigma^R$ given by

$$\delta \;\mapsto\; \lambda r{:}R.\ \exists e{:}Q.\ r < je \wedge \delta e \quad \text{and} \quad \lambda d{:}Q.\ \Delta(jd) \;\leftarrow\!\shortmid\; \Delta$$

restrict to a bijection between rounded lower $\delta : \Sigma^Q$ and $\Delta : \Sigma^R$. Hence the two constructions of $\underline{R}$ agree. Those for $\overline{R}$ do so too, with

$$\upsilon \;\mapsto\; \lambda r{:}R.\ \exists t{:}Q.\ jt < r \wedge \upsilon t \quad \text{and} \quad \lambda u{:}Q.\ \Upsilon(ju) \;\leftarrow\!\shortmid\; \Upsilon.$$

Proof Recall from Lemmas 6.12f that $jd <_R je \Leftrightarrow d <_Q e$ and $(\delta, \upsilon) <_R ju \Leftrightarrow \upsilon u$. Then

$$\Delta \;\mapsto\; \delta \;\mapsto\; \lambda r.\,\exists e{:}Q.\ r < je \wedge \Delta(je) \;=\; \lambda r.\,\exists s{:}R.\ r < s \wedge \Delta(s) \;\equiv\; \hat{\Delta}$$

by Theorem 9.2, and $\delta \;\mapsto\; \Delta \;\mapsto\; \lambda d.\,\exists e.\ jd <_R je \wedge \delta e \;\equiv\; \lambda d.\,\exists e.\ d <_Q e \wedge \delta e \;\equiv\; \hat{\delta}$. □

Proposition 9.5 Under this correspondence, the rounded pseudo-cut (Δ, Υ) of R is respectively disjoint, bounded or located iff the pseudo-cut (δ, υ) of Q has the same property.

Proof For disjointness, using Theorem 9.2,

$$\begin{aligned}
\nabla(\Delta, \Upsilon) \quad &\equiv \quad \exists rs{:}R.\ \Delta r \wedge (s < r) \wedge \Upsilon s \\
&\Leftrightarrow \quad \exists du{:}Q.\ \Delta(jd) \wedge (u < d) \wedge \Upsilon(ju) \\
&\Leftrightarrow \quad \exists du{:}Q.\ \delta d \wedge (u < d) \wedge \upsilon u \;\equiv\; \nabla(\delta, \upsilon),
\end{aligned}$$

and the argument for B and boundedness just omits $s < r$ and $u < d$.

If (δ, υ) is located then so is (Δ, Υ), because, since $jd <_R jt \Leftrightarrow e <_Q t \Rightarrow \delta e \vee \delta t$,

$$r < s \;\Rightarrow\; \exists et{:}Q.\ r < je < jt < s \;\Rightarrow\; (\exists e.\ r < je \wedge \delta e) \vee (\exists t.\ jt < s \wedge \upsilon t) \;\equiv\; \Delta r \vee \Upsilon s.$$

Conversely, if (Δ, Υ) is located then so is (δ, υ), because

$$d <_Q u \iff jd <_R ju \;\Rightarrow\; \Delta(jd) \vee \Upsilon(ju) \;\equiv\; \delta d \vee \upsilon u. \qquad □$$

Theorem 9.6 $(R, <)$ is a ***Dedekind-complete*** dense linear order without endpoints, in the sense that every cut $(\Delta, \Upsilon) : \Sigma^R \times \Sigma^R$ (possibly involving parameters) is of the form

$$\Delta \;=\; \lambda x.\,(x < a) \quad \text{and} \quad \Upsilon \;=\; \lambda y.\,(a < y)$$

for some unique $a : R$ (with the same parameters).

Proof The real number a is the cut (δ, υ) of Q given by the above correspondence. It satisfies

$$x < a < y \iff (\exists e{:}Q.\ x < je \wedge \delta e) \wedge (\exists t{:}Q.\ jt < y \wedge \upsilon t) \;\equiv\; \Delta x \wedge \Upsilon y. \qquad \square$$

Remark 9.7 In the previous section, the construction of R from Q involved the formulation of the nucleus $\mathcal{E}$. However, we cannot repeat this step with R in place of Q, because existential quantification over finite lists in R is not allowed in Notation 8.2.

Why on Earth not? This has nothing to do with rational or irrational numbers: so far the elements of Q are only nominally "rational", as we have not introduced any arithmetic yet. The point is topological: Q is overt discrete — essentially a "set" — and $\mathsf{List}(Q)$ is the overt discrete "hyperspace" of its overt compact subspaces (Theorem 5.21). On the other hand, R is not discrete but Hausdorff, and its overt compact subspaces can be infinite — bounded closed intervals, for example. In the hyperspace of such subspaces, the finite ones form a dense subspace, so there is no way to distinguish them from the infinite ones.

But this does not matter. We do not need to define another space $S \rightarrowtail \Sigma^R \times \Sigma^R$ from such an $\mathcal{E}$, as we have already done enough to prove completeness, so R itself serves for S. Then

$$R \rightarrowtail I\!R \rightarrowtail \underline{R} \times \overline{R} \rightarrowtail \Sigma^R \times \Sigma^R$$

is a composite of Σ-split inclusions. In fact, the splitting will be expressible in the same way as in Proposition 2.15, after we have proved compactness of $[d, u]$ (Proposition 14.7).

Two other important notions of completeness for the real line are considered in [J], namely the convergence of Cauchy sequences and the existence of suprema of nonempty bounded subsets. In any form of constructive analysis, the latter result requires additional hypotheses; in ASD these are that the subset be compact and overt.

A natural special case of the supremum is the maximum of a pair:

Proposition 9.8 R has binary join and meet with respect to $\leq$,

$$\begin{aligned} \mathsf{max}\big((\delta_1, \upsilon_1),\ (\delta_2, \upsilon_2)\big) \;&\equiv\; (\delta_1 \vee \delta_2,\ \upsilon_1 \wedge \upsilon_2) \\ \mathsf{min}\big((\delta_1, \upsilon_1),\ (\delta_2, \upsilon_2)\big) \;&\equiv\; (\delta_1 \wedge \delta_2,\ \upsilon_1 \vee \upsilon_2). \end{aligned}$$

satisfying $\mathsf{min}(a, b) \leq a, b \leq \mathsf{max}(a, b)$ and

$$\begin{array}{rcl@{\qquad}rcl} (a < x) \wedge (b < x) & \Leftrightarrow & \mathsf{max}(a,b) < x & (a > x) \wedge (b > x) & \Leftrightarrow & \mathsf{min}(a,b) > x \\ (a > x) \vee (b > x) & \Leftrightarrow & \mathsf{max}(a,b) > x & (a < x) \vee (b < x) & \Leftrightarrow & \mathsf{min}(a,b) < x \\ (a \leq x),\ (b \leq x) & \dashv\vdash & \mathsf{max}(a,b) \leq x & (a \geq x),\ (b \geq x) & \dashv\vdash & \mathsf{min}(a,b) \geq x. \end{array} \qquad \square$$

We have said that the two halves of a Dedekind cut play symmetrical roles. Adding an operation that makes this literally so, we have the beginnings of arithmetic:

Proposition 9.9 If $(Q, <)$ has an order-reversing automorphism $(-)$ then so does R, and this has a unique fixed point (0):

$$\ominus(\delta, \upsilon) \;\equiv\; \big(\lambda d.\,\upsilon(-d),\ \lambda u.\,\delta(-u)\big) \qquad 0 \;\equiv\; (\lambda d.\, d < -d,\ \lambda u.\ -u < u).$$

This agrees with Moore's interval arithmetic, since $\ominus(\delta_d, \upsilon_u) = (\delta_{-u}, \upsilon_{-d})$. Since it preserves roundedness, it defines an isomorphism $\underline{R} \cong \overline{R}$ (Definition 7.3), but not an automorphism of either of them. Since it also preserves disjointness, boundedness and locatedness, we have $\ominus : IR \cong IR$ and $- : R \cong R$.

From negation and maximum we may define the absolute value function,

$$|(\delta, \upsilon)| \;\equiv\; \mathsf{max}\big((\delta, \upsilon),\; -(\delta, \upsilon)\big) \;=\; \big(\lambda d.\, \delta d \vee \upsilon(-d),\; \lambda u.\, \upsilon u \wedge \delta(-u)\big),$$

so $|x| = x$ if $x \geq 0$ and $|x| = -x$ if $x \leq 0$. This satisfies $(x \neq 0) \equiv (\delta 0 \vee \upsilon 0) \Leftrightarrow |x| \neq 0$. ☐

We shall develop addition and multiplication in Sections 11 and 12, and we use Dedekind completeness to find the inverses of strictly monotone functions in Section 13, but now we turn to the Heine–Borel theorem.

10 Open, compact and overt intervals

In this section we prove that the closed interval is compact and overt, in the sense of Definition 4.14, namely that it admits both "quantifiers", $\forall$ and $\exists$. The constructions formalise in ASD some of the ideas that we met in the context of interval analysis in Section 2. We compute the universal quantifier $\forall x{:}[d,u].\, \phi x$ by splitting the interval into sufficiently but finitely many sub-intervals, and applying ϕ to each of them *à la* Moore. The existential quantifier, on the other hand, is related to the back-to-front (Kaucher) intervals that we saw in Remark 2.20.

First, of course, we have to *construct* the closed interval itself, and for the sake of symmetry the open one too, as objects of ASD. These are defined by means of terms of type Σ^R.

Definition 10.1 For $d \leq u : R$, the ***open*** and ***closed intervals*** $(d, u), [d, u] \subset R$ are
(a) the open subspace classified by $(\lambda x.\, d < x < u) \equiv (\upsilon_d \wedge \delta_u) : \Sigma^R$, and
(b) the closed subspace co-classified by $(\lambda x.\, x < d \vee u < x) \equiv (\delta_d \vee \upsilon_u) : \Sigma^R$ respectively.
More generally, for any rounded pair (δ, υ),
(c) $(\upsilon, \delta) \subset R$ is the open subspace classified by $\upsilon \wedge \delta : \Sigma^R$, and
(d) $[\delta, \upsilon] \subset R$ is the closed subspace co-classified by $\delta \vee \upsilon : \Sigma^R$,
where we expect the pseudo-cut (δ, υ) to be located (maybe overlapping or back-to-front) in the first case and disjoint in the second.

Remark 10.2 The ASD subspaces that are defined by these terms were constructed in Remarks 5.8ff. However, as we have not allowed dependent types in the calculus, we must temporarily restrict attention to intervals with constant endpoints, so d, u, δ and υ cannot have parameters. In other words, we are essentially just dealing with the *unit* interval $\mathbb{I} \equiv [0, 1]$, where $d \equiv 0$, $u \equiv 1$. But this is no real handicap, since (when we have some arithmetic) the general interval $[d, u]$ with endpoints will be the direct image of $[0, 1]$ under the function $t \mapsto d(1-t) + ut$. This actually generalises to all rounded bounded (δ, υ). The outcome is that the formulae that we obtain in the end remain valid with parametric endpoints [J].

Even if we had a theory of dependent types and tried to use it to prove the results in this section, we would still have to demonstrate that the formulae that *look* like nuclei or quantifiers *actually* correspond to (dependent) subspaces. That would be at least as difficult as the approach that we take here.

On the other hand, *terms* with parameters have been familiar for centuries, and translate more directly into programs. In [J] we will use modal operators (quantifiers over subspaces) derived from f as a tool for solving the equation $f(x) = 0$, naturally involving the same parameters as the function. These *operators* are (Scott) continuous in the parameters, which means that we may still work with them in the vicinity of singularities of the equation, where the *sets* or *types* of zeroes are discontinuous.

Proposition 10.3 Any open subspace $U \subset R$ classified by $\theta : \Sigma^R$ is overt. This applies in particular to the open intervals (d,u) and (υ,δ) in Definition 10.1. The existential quantifiers are defined for $\phi : \Sigma^R$ by

$$\begin{array}{rcl} \exists x{:}U.\ \phi x & \equiv & \exists x{:}R.\ \theta x \wedge \phi x \\ \exists x{:}(d,u).\ \phi x & \equiv & \exists x{:}R.\ (d < x < u) \wedge \phi x \\ \exists x{:}(\upsilon,\delta).\ \phi x & \equiv & \exists x{:}R.\ \upsilon x \wedge \delta x \wedge \phi x, \end{array}$$

in which $\exists x : R$ was defined in Theorem 9.2 and is the same as $\exists x : Q$.

Proof We may deduce in both directions:

$$\begin{array}{llll} \Gamma,\ x : U & \vdash & \phi x \ \Rightarrow\ \sigma & \\ \Gamma,\ x : R,\ \theta x \Leftrightarrow \top & \vdash & \phi x \ \Rightarrow\ \sigma & \\ \Gamma,\ x : R & \vdash & \theta x\ \wedge\ \phi x \ \Rightarrow\ \sigma & \text{Axiom 4.7} \\ \Gamma & \vdash & (\exists x{:}R.\ \theta x \wedge \phi x)\ \Rightarrow\ \sigma & R \text{ overt} \end{array}$$

so $\exists x{:}R.\ \theta x \wedge \phi x$ satisfies the defining property of $\exists x{:}U.\ \phi x$. □

Remark 10.4 The main task of this section is to show that, for $d \le u$, the closed interval $[d,u]$ is overt and compact. The quantifiers $\exists$ and $\forall$ will be given by

$$\exists x{:}[d,u].\ \phi x \ \Leftrightarrow\ \mathcal{E}_1\Phi(\delta_u,\upsilon_d)\ \Leftrightarrow\ \mathcal{E}\Phi(\delta_u,\upsilon_d) \quad \text{and} \quad \forall x{:}[d,u].\ \phi x \ \Leftrightarrow\ \mathcal{E}\Phi(\delta_d,\upsilon_u),$$

where Lemma 8.10 allows us to interchange $\mathcal{E}$ with $\mathcal{E}_1$ in the existential quantifier, because the argument (δ_u,υ_d) is located.

Here $\Phi : \Sigma^Q \times \Sigma^Q \to \Sigma$ is *any* extension of $\phi : R \to \Sigma$. As in Remark 2.5, the canonical one is $\Phi \equiv I\phi$, but the Moore–Kaucher interpretation of the arithmetical operations may be used to find another extension in a more practical way. In any case, the notation is simpler if we put things the other way round, working primarily with Φ, so that $\phi x \equiv \Phi(ix) \equiv \Phi(\delta_x,\upsilon_x)$.

We first show that the *closed* interval, like the open one, is *overt*, by a generalisation of Theorem 9.2. Since $\mathcal{E}_1$ preserves $\bot$, this may also be seen as an application of Lemma 5.13 to the lattice $\underline{R} \times \overline{R} \downarrow (\delta_u,\upsilon_d)$ of rounded pseudo-cuts with $(\delta,\upsilon) \le (\delta_u,\upsilon_d)$. Recall that $L \downarrow \alpha$ means $\{\beta : L \mid \beta \leqslant \alpha\}$.

Proposition 10.5 The closed interval $[d,u]$ is ***overt***, with $\exists x{:}[d,u].\ \Phi(ix)\ \equiv\ \mathcal{E}_1\Phi(\delta_{\boldsymbol{u}},\upsilon_{\boldsymbol{d}})$.

$$\frac{\Gamma,\ x : R,\ d \le x \le u\ \vdash\ \phi x\ \Rightarrow\ \sigma}{\Gamma\ \vdash\ \exists x{:}[d,u].\ \phi x\ \Rightarrow\ \sigma}$$

Notice that, as in Remark 2.20, u and d are the "wrong" way round.

Proof This is similar to Proposition 10.3. By Lemma 9.1, the top line of the rule is

$$\Gamma,\ x,d,u : R,\ d \le x \le u\ \vdash\ \phi x\ \equiv\ \exists et.\ e < x < t \wedge \Phi(\delta_e,\upsilon_t)\ \Rightarrow\ \sigma,$$

which, by Definition 4.14 (for $\exists$) and Axiom 4.7, is equivalent to

$$\Gamma,\ x,d,u,e,t : R,\ d \le x \le u,\ e < x < t\ \vdash\ \Phi(\delta_e,\upsilon_t)\ \Rightarrow\ \sigma.$$

The inequalities on the left of the $\vdash$ are equivalent to

$$d \le u, \quad \mathsf{max}(d,e) \ \le\ x \ \le\ \mathsf{min}(t,u), \quad e < t, \quad e < u, \quad d < t,$$

the last three of which can be moved back across the $\vdash$, using Definition 4.14 and Axiom 4.7 again, whilst x is now redundant. Hence

$$\Gamma,\ d,u : R,\ d \le u\ \vdash\ \mathcal{E}_1\Phi(\delta_u, \upsilon_d)\ \equiv\ \exists e < t.\, e < u \wedge d < t \wedge \Phi(\delta_e, \upsilon_t)\ \Rightarrow\ \sigma,$$

which is the bottom line of the rule. $\square$

Exercise 10.6 Show that $\exists x{:}[d,u].\ \phi x \implies \phi d \vee (\exists x{:}R.\ d < x < u \wedge \phi x) \vee \phi u.$ $\square$

Now we come to the principal result of this paper, that the closed interval $[d,u]$ is compact. The proof may be considered as another example of Lemma 5.13, with

$$[d,u]\ \subset\ \{(\delta,\upsilon) : IR \mid \delta_d \leqslant \delta\ \&\ \upsilon_u \leqslant \upsilon\}.$$

Theorem 10.7 The closed interval $[d,u] \subset R$ is ***compact*** in the sense of Definition 4.14:

$$\frac{\Gamma,\ x : [d,u]\ \vdash\ \sigma\ \Rightarrow\ \phi x\ \equiv\ \Phi(ix)}{\Gamma\ \vdash\ \sigma\ \Rightarrow\ \forall x{:}[d,u].\ \phi x\ \equiv\ I\phi(\delta_d, \upsilon_u)\ \equiv\ \mathcal{E}\Phi(\delta_d, \upsilon_u)}$$

Proof In the upward direction, let $(\delta,\upsilon) \equiv (\delta_x, \upsilon_x) \geqslant (\delta_d, \upsilon_u)$ be the cut corresponding to $x : R$. Then

$$\Gamma\ \vdash\ \sigma\ \Rightarrow\ I\phi(\delta_d, \upsilon_u)\ \Rightarrow\ I\phi(\delta,\upsilon)\ \equiv\ I\phi(ix)\ \Leftrightarrow\ \phi x.$$

For the converse, first let $\psi \equiv \lambda x.\,(x > u \vee x < d) : \Sigma^R$ and $\Psi \equiv \lambda\alpha\beta.\,\alpha u \vee \beta d : \Sigma^{\Sigma^Q \times \Sigma^Q}$, so $\psi = \Sigma^i\Psi$ co-classifies $[d,u]$ (Definition 10.1). Using Axiom 4.7, the top line says

$$\Gamma,\ x : R\ \vdash\ \phi x \vee \psi x\ \equiv\ \phi x \vee (x < d) \vee (x > u)\ \Leftrightarrow\ \top.$$

This captures universal quantification as a *judgement* (Definition 4.4). By λ-abstraction,

$$\Gamma\ \vdash\ \Sigma^i(\Phi \vee \Psi)\ =\ \Sigma^i\Phi \vee \Sigma^i\Psi\ =\ \phi \vee \psi\ =\ \top\ =\ \Sigma^i\top : \Sigma^R,$$

so if we apply I to this, where $\mathcal{E} = I \cdot \Sigma^i$, and the result to (δ_d, υ_u), we have

$$\mathcal{E}(\Phi \vee \Psi)(\delta_d, \upsilon_u)\ \Leftrightarrow\ \mathcal{E}\top(\delta_d, \upsilon_u)\ \Leftrightarrow\ \top$$

by Lemma 8.5. In the expansion of $\mathcal{E}(\Phi \vee \Psi)(\delta_d, \upsilon_u)$, the same Lemma uses conjuncts like

$$(\Phi \vee \Psi)(\delta_{q_k}, \upsilon_{q_{k+1}})\ \Leftrightarrow\ \Phi(\delta_{q_k}, \upsilon_{q_{k+1}})\ \vee\ (u < q_k \vee q_{k+1} < d)\ \Leftrightarrow\ \Phi(\delta_{q_k}, \upsilon_{q_{k+1}})\ \vee\ \bot.$$

Hence $$\top\ \Leftrightarrow\ \mathcal{E}\Phi(\delta_d, \upsilon_u)\ \Leftrightarrow\ \mathcal{E}(\Phi \vee \Psi)(\delta_d, \upsilon_u).$$ $\square$

Corollary 10.8 Re-substituting Definition 8.2 of $\mathcal{E}$,

$$\begin{array}{lcl}
\forall x{:}[d,u].\ \Phi(ix) & \Leftrightarrow & \mathcal{E}\Phi(\delta_d, \upsilon_u) \\
 & \Leftrightarrow & \exists d < \cdots < u.\ \bigwedge_{k=0}^{n-1} \Phi(\delta_{q_k},\ \upsilon_{q_{k+1}}) \\
 & \Leftarrow & \Phi(\delta_d, \upsilon_u) \\
\Phi(ix)\ \equiv\ \Phi(\delta_x, \upsilon_x) & \Rightarrow & \exists d < x < u.\ \Phi(\delta_d, \upsilon_u) \\
 & \Rightarrow & \exists d < x < u.\ \forall y{:}[d,u].\ \Phi(iy)\ \Rightarrow\ \Phi(ix) \\
\exists x{:}[d,u].\ \Phi(ix) & \Leftrightarrow & \mathcal{E}_1\Phi(\delta_u, \upsilon_d) \\
 & \Leftrightarrow & \exists et.\ t < u\ \wedge\ d < e\ \wedge\ \Phi(\delta_t, \upsilon_e) \\
 & \Rightarrow & \Phi(\delta_u, \upsilon_d),
\end{array}$$

using Notation 7.1, Lemmas 8.3, 8.5 and 9.1 and Proposition 10.5. $\square$

Warning 10.9 In view of the fact that both quantifiers are given by almost the same formula, the hypothesis $d \leq u$ must be important. If we do not know this, we can still obtain the bounded *universal* quantifier as

$$\big(\forall d \leq x \leq u.\, \Phi(ix)\big) \;\equiv\; (d > u) \;\vee\; \mathcal{E}\Phi(\delta_d, \upsilon_u),$$

but for the *existential* quantifier we have to write

$$\Phi : \Sigma^{\Sigma^Q \times \Sigma^Q},\; d \leq u : R \;\vdash\; \exists d \leq x \leq u.\, \Phi(ix) \;\equiv\; \mathcal{E}\Phi(\delta_u, \upsilon_d).$$

This is because $(\exists d \leq x \leq u.\, \top) \Leftrightarrow \top \;\dashv\vdash\; (d \leq u) \;\dashv\vdash\; (d > u) \Leftrightarrow \bot$, making a positive statement equivalent to a negative one. But that would make the predicate $(d > u)$ decidable, which it can only be if we had already *assumed* that it was either true or (in fact) false.

Notice that the intersection of two overt subspaces need not be overt, even in the simple case of $[d,u] \cap [e,t]$. This is because we need to be able to decide whether they are disjoint ($t < d \vee u < e$) or overlap ($t \geq d$ & $u \geq e$). We shall see more examples of this in [J].

While we're looking at awkward cases, recall from Theorem 9.2 that

$$\exists x{:}R.\; \Phi(ix) \iff \mathcal{E}\Phi(\top, \top),$$

and it's easy to check that $\exists x \leq u.\, \Phi(ix) \;\Leftrightarrow\; \mathcal{E}\Phi(\delta_u, \top)$ and $\exists x \geq d.\, \Phi(ix) \;\Leftrightarrow\; \mathcal{E}\Phi(\top, \upsilon_d)$.

So what happens if we let one or both of the arguments of $\mathcal{E}\Phi$ be $\bot$? Does this give a meaning to $\phi(\infty) \equiv \mathsf{lim}_{x\to\infty}\phi x$ or to $\forall{-\infty} \leq x \leq \infty.\, \phi x$? Unfortunately not: putting δ or $\upsilon \equiv \bot$ in Notation 8.2 just gives $\bot$. $\square$

Remark 10.10 You are probably wondering where the "finite open sub-cover" has gone. Essentially, it was absorbed into the axioms, in the form of general Scott continuity (Axiom 4.21).

Compactness is often said to be a generalisation of finiteness. We dissent from this. The finiteness is a side-effect of Scott continuity. The essence of *compactness* lies, not in that the cover be finite, but in what it says about the notion of *covering*. Recall how the *statement* of equality became an (in)equality *predicate* in a discrete or Hausdorff space (Remark 4.6). Similarly, compactness promotes the *judgement* (Definition 4.4),

$$\ldots,\; x : [d,u] \;\vdash\; \phi x \Leftrightarrow \top,$$

that an open subspace covers into a *predicate* (Axiom 4.2),

$$\forall x{:}[d,u].\; \phi x \;\equiv\; I\phi(\delta_d, \upsilon_u) \;\Leftrightarrow\; \top.$$

Now, $\lambda\phi.\, \forall x{:}[d,u].\, \phi x$ is a term in the ASD λ-calculus of type Σ^{Σ^R}, and, like all such terms, it preserves directed joins. So if $\phi_\ell x$ is directed with respect to $\ell : D$ then

$$\forall x{:}[d,u].\; \bigvee^{\uparrow}_{\ell} \phi_\ell x \iff \bigvee^{\uparrow}_{\ell} \forall x{:}[d,u].\; \phi_\ell x.$$

Restating this for general joins, we have

$$\forall x{:}[d,u].\;\; \exists n{:}N.\;\; \phi_n x \iff \exists \ell{:}\mathsf{Fin}(N).\;\; \forall x{:}[d,u].\; \exists n \in \ell.\; \phi_n x,$$

where ℓ provides the finite open sub-cover.

Now, it is a little surprising that the Heine–Borel property is traditionally expressed using *finite* open sub-covers, when a very common situation in analysis is that the cover is indexed by the real numbers, which are directed. In this case, *just one* member of the family covers, and we have *uniform* values, for example

$$\forall x{:}[d,u].\ \exists\delta>0.\ \phi_\delta x \iff \exists\delta>0.\ \forall x{:}[d,u].\ \phi_\delta x,$$

so long as ϕ is contravariant in the sense that $\delta\le\epsilon\ \vdash\ \phi_\epsilon\ \Rightarrow\ \phi_\delta$. $\square$

Warning 10.11 Our "$\exists\ell:\mathsf{Fin}(N)$" and $\exists\delta>0$ do not have the same strength as they do in other constructive systems. The computational implementation only *produces* ℓ or δ in the case where the predicate (is provably true and) has at worst variables of type $\mathbb{N}$ (Remark 3.5). In particular, if the predicate ϕ (and maybe even the bounds d and u) are expressions with real-valued parameters, neither δ nor even the length of the list ℓ can be extracted as functions of these parameters.

Lemma 10.12 The *open* interval $(0,1)$ is not compact.

Proof By Definition 6.10 for $0<x$,

$$x:(0,1)\ \vdash\ \top\ \Leftrightarrow\ \exists q{:}Q.\ 0<q<x,$$

where $\exists$ is directed, so if $(0,1)$ were compact we could interchange the quantifiers, but

$$\top\iff\forall x{:}(0,1).\ \exists q>0.\ q<x\iff\exists q>0.\ \forall x{:}(0,1).\ q<x\iff\bot. \qquad\square$$

11 Arithmetic

Real arithmetic is a case where familiarity breeds contempt. Most of the issues that we need to consider are extremely well known as *folklore*, but we do not know of a published account that considers them *all* and gives the details in a coherent form that is suitable for us. See [TvD88, p260] for a metrical approach involving 2^{-n}, but our paper is about Dedekind cuts and intervals, which are fundamentally *order*-theoretic ideas.

Dedekind himself considered addition and square roots but not multiplication [Ded72]. Even in the classical situation, the proof that the *positive and negative* reals form a commutative ring or field involves either an explosion of cases or considerable ingenuity. One ingenious solution is John Conway's definition of multiplication [Con76, pp.19–20], but we shall not use it.

Working with intervals instead of numbers makes this problem worse, whilst its usual solution — case splitting — is unacceptable, at least for R, in constructive analysis [TvD88, §5.6], topos theory [FH79] [Joh77, §6.6] or ASD.

On the positive side, general continuity will relieve us of the burden of checking the arithmetic laws, and in particular the case analysis in R. However, (Scott) continuity is only applicable after we have verified an order-theoretic property of *discrete* rational arithmetic (*cf.* Proposition 7.13). In Q we can and do use case analysis.

For reasons of space we omit the multiplication sign ($\times$) in the usual way. Juxtaposition may therefore denote either multiplication or function application, but our alphabetic conventions distinguish them: ab with two italic letters is multiplication, whilst ϕa with a Greek letter is application. Despite the weight of tradition, we use p instead of ϵ for a small positive tolerance, saving (δ,υ) and (ϵ,τ) for the arguments of an arithmetic operation on cuts, with result (α,σ).

As usual, multiplication and λ-application bind more tightly than $<$, which is tighter than $\wedge$.

Axiom 11.1 Q is a discrete, densely linearly ordered (Definition 6.1) commutative ring.

There is a unique "inclusion" map $\mathbb{Z} \subset Q$ that preserves and reflects $=$, $<$, 0, 1, $+$, $-$ and $\times$.

Axiom 11.2 Q obeys the ***Archimedean principle***, for $p, q : Q$,

$$q > 0 \implies \exists n{:}\mathbb{Z}.\ q(n-1) < p < q(n+1).$$

In view of Proposition 6.5, this cannot be a property of the *order* $(Q, <)$ alone: it says how $\mathbb{Z}$ (*i.e.* the additive subgroup generated by q) lies within the order. We isolate our use of this principle in Propositions 11.15 and 12.7, in the hope that some future development may avoid it altogether (Remark 12.13).

Remark 11.3 We do not ask that Q be a field, with (exact) division, because we have in mind the ways in which it might be represented computationally: the axioms ought at least to allow the dyadic and decimal rationals as examples. We could weaken the properties of addition and multiplication on Q still further, since (in Example 11.9 and Proposition 12.3) we shall only use them in the form of the four ternary relations

$$a < d + e \qquad u + t < z \qquad a < d \times u \qquad u \times t < z,$$

whose axiomatisation we leave to the interested reader. Such a person may be an interval analyst or programmer who wants to use ASD to obtain mathematically provable results from the ***floating point arithmetic*** that is built in to computer hardware [Ste85].

Whilst the likely applications have it, we do not actually even *need* to assert division by 2 as an axiom:

Lemma 11.4 There is an ***approximate half***, $h : Q$, with $0 < h + h < 1$.

Proof Let $a, h : Q$ with $0 < a < 1$ and $0 < h < \mathsf{min}(a, 1-a)$. □

Since the operations of arithmetic take two arguments, and its equations involve (up to) three variables, our first task is to define the objects $R^{\mathbf{2}}$ and $R^{\mathbf{3}}$. By $R^{\mathbf{n}}$ here we simply mean the product of 2, 3, … copies.

Proposition 11.5 The product spaces $R^{\mathbf{n}}$ exist, and they are overt, with $\exists_{R^{\mathbf{n}}} = \exists_{Q^{\mathbf{n}}}$.

Proof Axiom 4.1 gives product types, and by Lemma 9.1 the quantifiers are given by

$$\begin{aligned} \exists(x,y){:}R^{\mathbf{2}}.\ \phi(x,y) \quad &\equiv \quad \exists x{:}R.\ \exists y{:}R.\ \phi(x,y) \\ &\Leftrightarrow \quad \exists x{:}Q.\ \exists y{:}Q.\ \phi(x,y) \ \equiv\ \exists(x,y){:}Q^{\mathbf{2}}.\ \phi(x,y). \end{aligned}$$ □

Remark 11.6 Now we have a diagram that combines those in Remarks 2.13 and 6.9,

$$\begin{tikzcd} Q^{\mathbf{2}} \arrow[r] \arrow[d, "*"'] & R^{\mathbf{2}} \arrow[r, tail] \arrow[d, dotted] & (\Sigma^Q \times \Sigma^Q)^{\mathbf{2}} \arrow[r, shift left] \arrow[r, shift right] \arrow[d, "\circledast"] & \Sigma^{S \times \mathbf{2}} \\ Q \arrow[r] & R \arrow[r, tail] & \Sigma^Q \times \Sigma^Q \arrow[r, shift left] \arrow[r, shift right] & \Sigma^S \end{tikzcd}$$

where $\rightarrowtail$ denotes an equaliser and $S \equiv Q + Q + \mathbf{1} + \mathbf{1} + Q \times Q + Q \times Q$.

This means that a term of type $R^{\mathbf{2}}$ is a pair of solutions of equations, defining a pair of cuts or a "Dedekind cross-hair", *i.e.* a division of the rational plane into quadrants. $R^{\mathbf{2}}$ is, in fact, a Σ-split subspace [B], though the splitting I_2 is not simply the square of that for R. But that does not matter, as we shall not need it anyway.

Exercise 11.7 By analogy with Proposition 2.15, find $I_2 : \Sigma^{R\times R} \rightarrowtail \Sigma^{\Sigma^Q\times\Sigma^Q\times\Sigma^Q\times\Sigma^Q}$.

Remark 11.8 The plan, for each arithmetic operation $*$ ($+$ or $\times$), is

(a) first to define some operation $\circledast$ on $\Sigma^Q \times \Sigma^Q$ that extends the given operation $*$ on the rationals, in the sense that the rectangle above commutes, *i.e.* for each $q, r : Q$,

$$(\delta_q, \upsilon_q) \circledast (\delta_r, \upsilon_r) \;=\; (\delta_{q*r}, \upsilon_{q*r}) : \Sigma^Q \times \Sigma^Q;$$

(b) then to show that $(\delta,\upsilon) \circledast (\epsilon,\tau)$ is a Dedekind cut whenever (δ,υ) and (ϵ,τ) are (so the composites $R^{\mathbf{2}} \rightrightarrows \Sigma^S$ above are equal), thereby filling in the dotted arrow above;

(c) and finally to verify the usual laws of arithmetic for the extended operations.

Classically, Ramon Moore has already done the first of these tasks, but we need to take account of the more general intervals that exist constructively (Remarks 2.3f). Recall from Propositions 7.12f that the intervals with (rational) endpoints provide a *basis* in the sense of domain theory, and that we may extend a predicate $\phi(d,u)$ to $\Phi(\delta,\upsilon)$ such that $\Phi(\delta_d,\upsilon_u) \Leftrightarrow \phi(d,u)$ so long as ϕ is rounded, in the sense that

$$\phi(e,t) \iff \exists du.\, (d<e) \wedge (t<u) \wedge \phi(d,u).$$

This condition may easily be generalised to an n-ary operation, but that would lead to another explosion of symbols and cases. It is therefore wise to consider the arguments and the bounds one at a time, which is enough.

Example 11.9 Consider the bounds a, z on addition of a fixed rational interval $[e,t]$, so

$$\phi(d,u) \quad\equiv\quad a < [d,u] \oplus [e,t] < z \quad\equiv\quad a < d+e \;\wedge\; u+t<z.$$

Then this condition requires $a < d+e \Leftrightarrow \exists d'.\, (d'<d) \wedge (a < d'+e)$, for which we choose d' with $a-e<d'<d$, using subtraction and interpolation, and similarly for the upper bound. □

Notation 11.10 For $(\delta,\upsilon), (\epsilon,\tau) : \Sigma^Q \times \Sigma^Q$, let

$$(\delta,\upsilon) \oplus (\epsilon,\tau) \quad\equiv\quad \big(\lambda a.\, \exists de.\, (a<d+e) \wedge \delta d \wedge \epsilon e, \quad \lambda z.\, \exists ut.\, (u+t<z) \wedge \upsilon u \wedge \tau t\big).$$

Remark 11.11 It's worth spelling out the condition in Proposition 7.13 once more in interval notation. Write x for (the name of) an interval with rational endpoints $[\underline{\mathsf{x}}, \overline{\mathsf{x}}]$, and

$$\mathsf{x} \Subset \mathsf{y} \quad\equiv\quad [\underline{\mathsf{x}}, \overline{\mathsf{x}}] \subset (\underline{\mathsf{y}}, \overline{\mathsf{y}}) \quad\equiv\quad \underline{\mathsf{y}} < \underline{\mathsf{x}} \;\wedge\; \overline{\mathsf{x}} < \overline{\mathsf{y}}$$

for the *strict* containment relation that relates the *basic compact* subspace $[\underline{\mathsf{x}}, \overline{\mathsf{x}}]$ to the *basic open* one $(\underline{\mathsf{y}}, \overline{\mathsf{y}})$, *cf.* [K]. (Beware that $\Subset$ corresponds to $\gg$ in the bounded interval predomain IR, which is a continuous dcpo.) Then the condition for $\circledast$ to be well defined is

$$\mathsf{x} \circledast \mathsf{y} \Subset \mathsf{w} \iff \exists \mathsf{x}'\mathsf{y}'.\; \mathsf{x} \Subset \mathsf{x}' \;\wedge\; \mathsf{y} \Subset \mathsf{y}' \;\wedge\; \mathsf{x}' \circledast \mathsf{y}' \Subset \mathsf{w}.$$

Note that $\exists$ here ranges over the *names* of the intervals, *i.e.* over $\{(d,u) \mid d \le u\} \subset Q^{\mathbf{2}}$, which is a discrete space, not over the domain IR, since $\Subset$ is not defined there as an ASD predicate in the sense of Axiom 4.2, but see [G, § 11].

Since it is only required to compute $\mathsf{x} \circledast \mathsf{y}$ to the limited precision w, it is standard programming practice to use less precise versions x' and y' of x and y in this way, thereby saving space and hence time.

Using roundedness, Proposition 7.12 extends Moore's operations to general intervals by

$$(\delta,\upsilon) \circledast (\epsilon,\tau) \;\equiv\; (\alpha,\sigma), \quad \text{where} \quad \alpha\underline{\mathsf{w}} \wedge \sigma\overline{\mathsf{w}} \;\equiv\; \exists \mathsf{xy}.\, \delta\underline{\mathsf{x}} \wedge \upsilon\overline{\mathsf{x}} \wedge \epsilon\underline{\mathsf{y}} \wedge \tau\overline{\mathsf{y}} \wedge \mathsf{x} \circledast \mathsf{y} \Subset \mathsf{w}.$$

Turning to stage (b) of the plan in Remark 11.8, we have to check that each formula that we introduce for an operation actually gives rounded, bounded, disjoint and located pairs (α,σ). Roundedness and boundedness follow immediately from the *form* of the definition that we have just given. We also get disjointness for free:

Lemma 11.12 Let $\alpha,\sigma : R^{\mathbf{n}} \rightrightarrows \Sigma^Q$ be "disjoint" on $Q^{\mathbf{n}}$, in the sense that

$$a,z:Q,\ \vec{q}:Q^{\mathbf{n}} \;\vdash\; \alpha\vec{q}a \wedge (z<a) \wedge \sigma\vec{q}z \;\Rightarrow\; \bot.$$

Then they are also disjoint on $R^{\mathbf{n}}$ in the same sense, *cf.* Proposition 9.5.

Proof By Proposition 11.5, $\alpha\vec{x}a \wedge (z<a) \wedge \sigma\vec{x}z \;\Rightarrow\; \exists\vec{q}.\, \alpha\vec{q}a \wedge (z<a) \wedge \sigma\vec{q}z \;\Rightarrow\; \bot$. ☐

Lemma 11.13 Addition (Notation 11.10) takes intervals (*i.e.* rounded, bounded and disjoint pseudo-cuts) to intervals, $\oplus : IR \times IR \to IR$.

Proof $(\delta,\upsilon) \oplus (\epsilon,\tau)$ is *rounded* by transitivity and interpolation in $(Q,<)$, keeping the same d,e,u,t. It is *bounded* because (δ,υ) and (ϵ,τ) are bounded, and by extrapolation in $(Q,<)$. It is *disjoint* by the previous Lemma, but more directly,

$$\begin{array}{rcl} \nabla\big((\delta,\upsilon)\oplus(\epsilon,\tau)\big) & \equiv & \exists azduet.\ \delta d \wedge \upsilon u \wedge \epsilon e \wedge \tau t \;\wedge\; (u+t<z<a<d+e) \\ & \Rightarrow & (\exists du.\ \delta d \wedge u<d \wedge \upsilon u) \;\vee\; (\exists et.\ \epsilon e \wedge t<e \wedge \tau t) \\ & \equiv & \nabla(\delta,\upsilon) \vee \nabla(\epsilon,\tau) \;\Leftrightarrow\; \bot \end{array}$$

since $u+t<d+e \Leftrightarrow (u-d)+(t-e)<0 \Rightarrow (u<d) \vee (t<e)$. ☐

Remark 11.14 *Locatedness* always seems to be the most difficult step, because we need to calculate the result arbitrarily closely. Recall that we defined this in a purely order-theoretic way in Definition 6.8, namely $d<u \Rightarrow \delta d \vee \upsilon u$.

However, to prove this property of a *sum*, we shall need a stronger hypothesis on the summ*ands*. This is actually the version of locatedness that most accounts of Dedekind cuts use, except that we allow the tolerance p to be any positive element of Q, not just 2^{-n} for $n:\mathbb{N}$.

It is not difficult to see the necessity of the stronger hypothesis for what we're trying to achieve, namely that Q is to be dense in a Dedekind-complete Abelian group R. In this situation, for any $x \equiv (\delta,\upsilon) : R$, there must be $d,u:Q$ with $x-\frac{1}{2}p<d<x<u<x+\frac{1}{2}p$. This is the (apparently essential) use of the Archimedean principle, to which we shall return in Remark 12.12.

Proposition 11.15 Any cut (δ,υ) is ***arithmetically located***, *i.e.* for $p:Q$,

$$p>0 \implies \exists et{:}Q.\ \delta e \wedge \upsilon t \wedge (0<t-e<p).$$

Conversely, if (δ,υ) is rounded and arithmetically located then it is bounded and located.

Proof By boundedness, $\delta d \wedge \upsilon u$ for some $d < u : Q$. Then the Archimedean principle gives some $k : \mathbb{N}$ with $u - d < khp$, where $0 < h + h < 1$ by Lemma 11.4. So $q_j \equiv d + jhp$ satisfy

$$q_0 \equiv d < q_1 < q_2 < \cdots < q_{k-1} < q_k > u,$$

to which we may apply Lemma 6.16, as (δ, υ) is order-located with δq_0 and υq_k. This gives $\delta e \wedge \upsilon t \wedge (t - e < p)$, where $e \equiv q_j$, $t \equiv q_{j+2}$ and $t - e = 2hp < p$, for some $0 \leq j \leq k - 2$.

Conversely, given $d < u$, let $e, t : Q$ with $t - e < u - d$ such that δe, υt, so (δ, υ) are bounded. By locatedness of $<$ (Lemma 6.2), $e < d \vee t < u$, so $\delta d \vee \upsilon u$ by roundedness. $\square$

Lemma 11.16 Addition takes cuts to cuts.

Proof Given $a < z$, we use *arithmetic* locatedness of one summand, (δ, υ), with $p \equiv z - a$, to give $d, u : Q$ such that δd, υu, $d < u$ and $(u - d) < (z - a)$. Since $a - d < z - u$, by interpolation there are $e, t : Q$ with $a - d < e < t < z - u$, so $a < d + e$ and $u + t < z$. By *order* locatedness of the other summand, $\epsilon e \vee \tau t$, so we have shown that

$$\exists deut{:}Q.\ \ (a < d + e) \wedge \delta d \wedge (\epsilon e \vee \tau t) \wedge (u + t < z) \wedge \upsilon u.$$

This implies that (α, σ) is order-located, *i.e.*

$$\big(\exists de.\,(a < d + e) \wedge \delta d \wedge \epsilon e\big) \ \vee\ \big(\exists ut.\,(u + t < z) \wedge \upsilon u \wedge \tau t\big). \qquad \square$$

Exercise 11.17 Show directly that
(a) $(\delta_q, \upsilon_q) \oplus (\delta, \upsilon)$ is order-located iff (δ, υ) is, for any $q : Q$; and
(b) if (δ, υ) and (ϵ, τ) are both arithmetically located then so is their sum. $\square$

In the final step (c) of Remark 11.8, topology plays to our advantage, particularly in ASD, where all definable maps are automatically continuous. Once we have thought of a definition for the extended operations, and verified that they have real results (*i.e.* cuts), the laws of arithmetic (both equations and inequalities) transfer automatically from $Q^{\mathbf{n}}$ to $R^{\mathbf{n}}$.

Lemma 11.18 For $f, g : R^{\mathbf{n}} \rightrightarrows R$,

$$\text{if}\quad \vec{q} : Q^{\mathbf{n}} \vdash f(\vec{q}) \leq g(\vec{q}) : R \quad\text{then}\quad \vec{x} : R^{\mathbf{n}} \vdash f(\vec{x}) \leq g(\vec{x}) : R.$$

Proof In traditional notation, since R is Hausdorff and f, g are continuous, the subspace

$$C \ \equiv\ \{\vec{x} : R^{\mathbf{n}} \mid f(\vec{x}) \leq g(\vec{x})\} \ \subset\ R^{\mathbf{n}}$$

on which the laws actually hold is closed; on the other hand, $Q^{\mathbf{n}} \subset R^{\mathbf{n}}$ is dense, whilst $Q^{\mathbf{n}} \subset C$ by hypothesis, so $C = R^{\mathbf{n}}$. In ASD, C is co-classified by θ, where

$$\theta(\vec{x}) \ \equiv\ f(\vec{x}) > g(\vec{x}), \quad \text{and} \quad \theta(\vec{x}) \ \Rightarrow\ \exists \vec{q}.\,\theta(\vec{q}) \ \Leftrightarrow\ \bot$$

by Proposition 11.5, so $\theta = \bot$. $\square$

Proposition 11.19 R is an ordered Abelian group. $\square$

12 Multiplication

We would like to develop multiplication in a way that follows addition as closely as possible. In that case, Example 11.9 showed that the extension of the Moore operation is well defined, and Lemma 11.16 that it preserves locatedness. The first of these tasks is complicated for multiplication by negation and intervals, including the back-to-front ones (Remark 2.20):

Lemma 12.1 ***Kaucher multiplication***[3] $[a, z] \equiv [d, u] \otimes [e, t]$ is defined by

$[d,u] \otimes [e,t]$		$d \le 0$, $u \le 0$	$d \le 0$, $u \ge 0$	$0 \le d$, $u \le 0$	$0 \le d$, $u \ge 0$
$0 \le e$	$0 \le t$	$[dt,\ ue]$	$[dt,\ ut]$	$[de,\ ue]$	$[de,\ ut]$
$0 \le e$	$t \le 0$	$[ut,\ ue]$	$[0, 0]$	$[q, p]$	$[de,\ dt]$
$e \le 0$	$0 \le t$	$[dt,\ de]$	$[p, q]$	$[0, 0]$	$[ue,\ ut]$
$e \le 0$	$t \le 0$	$[ut,\ de]$	$[ue,\ de]$	$[ut,\ dt]$	$[ue,\ dt]$

where $p \equiv \mathsf{min}(dt, ue) \le 0$ and $q \equiv \mathsf{max}(de, ut) \ge 0$. It agrees with ordinary multiplication in Q when $d = u$ and $e = t$, and with Moore's operation (Definition 2.3) when $d \le u$ and $e \le t$. Also, a is a monotone function of d and e and an antitone one of u and t, and *vice versa* for z. $\square$

Even though this is rather complicated, we can break it up into small pieces:

Lemma 12.2 The following functions are defined on the ascending reals, $\underline{R}$:
(a) constants;
(b) multiplication by $q \ge 0$: $\delta \mapsto \lambda a.\ \exists d.\ \delta d \wedge (a < dq)$;
(c) minimum, $\mathsf{min}(\delta, \epsilon) \equiv \lambda d.\ \delta d \wedge \epsilon d$;
(d) maximum, $\mathsf{max}(\delta, \epsilon) \equiv \lambda d.\ \delta d \vee \epsilon d$; and
(e) addition, $\delta \oplus \epsilon \equiv \lambda a.\ \exists de.\ \delta d \wedge \epsilon e \wedge a < d + e$.
There are similar functions on $\overline{R}$, and negation interchanges them. $\square$

Proposition 12.3 Kaucher multiplication provides a map

$$(\Sigma^Q \times \Sigma^Q) \times (\Sigma^Q \times \Sigma^Q) \longrightarrow (\Sigma^Q \times \Sigma^Q)$$

that extends multiplication on Q in the sense of Remark 11.5. It also preserves roundedness, boundedness and disjointness.

Proof Looking at pairs of entries in the table, we observe, for example, that a depends on d for fixed e, t and u in at worst a piecewise linear way. The pieces are either constant or multiplication by some $q \ge 0$ (or $q \le 0$ in the order-reversing cases), and they are combined using min or max. These functions are rounded on Q, or defined $\underline{R} \to \underline{R}$ *etc.*, by the previous lemma, and it's enough to check this componentwise.

Alternatively, we may follow Example 11.9 directly:

$$a < [d, u] \otimes [e, t] < z \quad \text{means} \quad a < 0,\ de,\ dt,\ ue,\ ut\ , \mathsf{min}(dt, ue),\ \mathsf{max}(de, ut) < z$$

in various cases, in each of which there is $d' < d$, $e' < e$, $t < t'$ or $u < u'$ with the same property.

[3]We computed the table ourselves with the help of Chris Stone, before we obtained a copy of [Kau80], and then checked that they agree.

Hence we have a map $(\underline{R} \times \overline{R}) \times (\underline{R} \times \overline{R}) \to (\underline{R} \times \overline{R})$, but $\underline{R}$ and $\overline{R}$ are retracts of Σ^Q, so we also have one $(\Sigma^Q \times \Sigma^Q) \times (\Sigma^Q \times \Sigma^Q) \to (\Sigma^Q \times \Sigma^Q)$ that preserves roundedness.

Since the Kaucher product of intervals with rational endpoints, *i.e.* as in the table, is finite, it preserves boundedness of general intervals. Similarly, since it agrees with Moore's operation for disjoint intervals, it also preserves disjointness of general intervals. Beware, however, that the product of a forward and a backward interval may be *either* way round, depending on the widths of the arguments. □

Exercise 12.4 Use the argument so far to show that R is a Q-module, *cf.* [Joh77, §6.6].

Remark 12.5 The final task is to show that products are located, *cf.* Lemma 11.16. Let's look at this from a programmer's perspective again. If we are asked to calculate xy to precision $0 < p < 1$, we need to decide how precisely to compute the factors x and y. For addition, $p/2$ is fine (*cf.* Exercise 11.17(b)), but for multiplication,

(a) if x and y are both small in magnitude, *i.e.* $|x|, |y| < 1$, then it suffices to find each of them to within p; but

(b) if x is large (legally, if $0 < M < |x|$, but we're thinking of the situation where M is in the millions), we need to find y correspondingly more precisely, to within p/M;

(c) similarly, if y is large then we need to know x more precisely.

Since we did not include division in the assumptions about Q, whereas we used subtraction in Example 11.9, we have a difficulty in the construction of multiplication even for *positive* real numbers. In any place-value representation, such as binary floating point, division may be performed as accurately as required, by first shifting the divisor by sufficiently many (n) places and then dividing to give an *integer* quotient. Proposition 12.7 is the same idea in abstract form; in the next section we shall show that this is enough to provide genuine division in R.

Lemma 12.6 Any linearly ordered ring Q is an ***integral domain***, admitting cancellation:

$$\text{if} \quad q > 0 \quad \text{then} \quad bq = cq \iff b = c \quad \text{and} \quad bq < cq \iff b < c. \qquad \square$$

Proposition 12.7 Q has ***approximate division***, for $a, q, z : Q$,

$$(a < z) \wedge (q > 0) \implies \exists m{:}Q.\ (a < mq < z).$$

Proof Either $0 < 4q < z - a$ or $0 < z - a < 8q$.

In the first case, we may apply the Archimedean principle directly with the given q:

$$\text{for some } k, k' : \mathbb{Z}, \quad q(k-1) < a < q(k+1) \quad \text{and} \quad q(k'-1) < z < q(k'+1).$$

Then $4q < z - a < q(k'+1) - q(k-1)$, so $k < k' - 2$ by the previous lemma. Hence $m \equiv k+1 < k'-1$, where $m : \mathbb{Z} \subset Q$, has the required property.

In the second case, the Archimedean principle for $8q$ (as p) and $z - a$ (as q) provides $k : \mathbb{Z}$ with $8q < 2n(z - a)$, and then $0 < k < 2^n$ for some $n : \mathbb{N}$. Let h satisfy $0 < h + h < 1$ (Lemma 11.4). Then

$$4q' \equiv 4qh^n < kh^n(z-a) < (2h)^n(z-a) < (z-a),$$

for which the first case gives $a < m'q' = m'qh^n < z$, so $m \equiv m'h^n$ is the required approximate quotient. □

Returning to the question of locatedness, recall that we needed to strengthen this notion in Proposition 11.15 in order to define addition. Curiously, we do *not* need to do this (or to apply the Archimedean principle) again for multiplication.

Lemma 12.8 Any arithmetically located positive cut $(\delta, \upsilon) : R$ is ***multiplicatively located***:

$$(0 < a < z) \wedge \delta 0 \implies \exists du{:}Q.\ (0 < d < u) \wedge \delta d \wedge \upsilon u \wedge (ua < dz),$$

where the last conjunct corresponds to $u - d < p$ in Proposition 11.15.

Proof By roundedness of δ, let $0 < r : Q$ with δr. By approximate division in Q, let $0 < p : Q$ with $0 < zp < r(z - a)$. By arithmetic locatedness, let $0 < d < u : Q$ with $\delta d \wedge \upsilon u \wedge (u - d < p)$. Since we have $\delta r \wedge \upsilon u$, disjointness of (δ, υ) gives $r < u$, so $z(u - d) < zp < r(z - a) < u(z - a)$. Hence $ua < zd$ as required. $\square$

Lemma 12.9 The product of an arithmetically located positive cut (δ, υ), *i.e.* such that $\delta 0$, with any cut (ϵ, τ) is another cut (α, σ), *cf.* Lemma 11.16.

Proof Suppose that $0 < a < z$. By multiplicative locatedness of (δ, υ), there are $0 < d < u$ with $\delta d \wedge \upsilon u \wedge (ua < dz)$. Then, using approximate division by du, there are e, t with

$$au < due < dut < zd.$$

So $a < de$, $e < t$ and $ut < z$ by Lemma 12.6, whilst $\epsilon e \vee \tau t$ by order-locatedness. Hence

$$(\exists de.\ a < de \wedge \delta d \wedge \epsilon e) \ \vee\ (\exists ut.\ ut < z \wedge \upsilon u \wedge \tau t) \ \equiv\ \alpha a \vee \sigma z.$$

More generally, given $a < z$, either

- $0 < z$, in which case there is some a' with $0, a \le \mathsf{max}(0, a) < a' < z$, so $\alpha a' \vee \sigma z$ by the foregoing argument, and hence $\alpha a \vee \sigma z$ since α is lower; or
- $a < 0$, where we apply the previous case to $-z < -a$, $\ominus(\epsilon, \tau)$ and $\ominus(\alpha, \sigma)$. $\square$

It only remains to prove locatedness of the product of two numbers that are both small. Note that we can bound a product *away* from zero iff we can do so for both factors, and in that case the previous result applies. The point of the fifth case below is therefore to constrain the product to be *near* to zero.

Lemma 12.10 The product of any two real numbers (cuts) $x, y : R$ is another cut.

Proof If $a < z$ then $a < 0 \vee 0 < z$, so $0 < m \equiv \mathsf{max}(z, -a)$ and

$$(x > 0) \ \vee\ (x < 0) \ \vee\ (y > 0) \ \vee\ (y < 0) \ \vee\ (|x| < 1 \ \wedge\ |y| < m).$$

It only remains to consider the last of these five cases, which is itself a disjunction because of the definition of $< \mathsf{max}$ (Proposition 9.8). Then essentially

$$|x| < 1 \ \wedge\ |y| < m \quad \implies \quad a < -|y| < xy \quad \vee \quad xy < |y| < z.$$

We need to explain these inequalities in terms of cuts $x \equiv (\delta, \upsilon)$ and $y \equiv (\epsilon, \tau)$:

$$\begin{array}{lcl} |x| < 1 & \equiv & -1 < x < +1 \ \equiv\ \delta d \wedge \upsilon u, \qquad \text{where} \quad d \equiv -1, \quad u \equiv +1 \\ a < -|y| & \equiv & \exists et.\ \epsilon e \wedge \tau t \ \wedge\ (a < e) \wedge (a < -t) \\ a < xy & \Leftarrow & \exists duet.\ \delta d \wedge \upsilon u \wedge \epsilon e \wedge \tau t \ \wedge\ a < \mathsf{min}(de, dt, ue, ut). \end{array}$$

$\square$

Proposition 12.11 R is an ordered commutative ring. $\square$

We shall use division in the next section to prove that R is Archimedean, but let us consider very briefly the necessity of that hypothesis on Q.

Remark 12.12 There are *Cauchy*-complete ordered fields with infinitesimals but, classically, any *Dedekind*-complete Abelian group must be Archimedean. This is because the sets

$$D \equiv \{d \mid \exists n{:}\mathbb{N}.\ d < n\} \quad \text{and} \quad U \equiv \{u \mid \forall n{:}\mathbb{N}.\ n < u\}$$

form a Dedekind cut that is located only in the weaker order-theoretic sense: since every $u - d$ is infinite, (D, U) is not arithmetically located.

The significance of the Archimedean property in Greek mathematics was recognised by Otto Stolz [Sto83]. He coined the name because, although the principle had been used "implicitly" by Eudoxus and Euclid, Archimedes had stated it as an Axiom. (He had also made far deeper use of it, in his *Method*, but Stolz was writing before the discovery of the most important Archimedes codex.)

Stolz also gave this argument that Dedekind completeness implies the Archimedean principle [*ibid.*, p. 511]. His result was disputed by his contemporaries, possibly because of its context in the debate at the time over the axiomatics of Euclid, but there is also a lot of extraneous discussion that makes one wonder whether he fully understood Dedekind completeness. However, we consider that his proof is valid, because it includes the two key points, namely the construction of the limit of an increasing sequence as the cut (D, U), and the problem with the value $(D, U) - 1$. The *constructive* least upper bound principle and arithmetic locatedness are also "implicit" in this paper.

Remark 12.13 What does this argument say about Conway's number system? Recall that it is a proper class. Although the class D above is equivalent as a left cut to the set $\mathbb{N}$, the class U cannot be expressed as a right cut, so $\{D \mid U\}$ is not a legitimate Conway Number — he calls it a *gap* [Con76, p. 37].

The argument also fails in ASD, for an analogous reason: U is not *definable* in the calculus as an open subspace. The point is really that the space D of *finite* numbers is *overt*, as it is given by an existential quantifier, or as the numbers for which repeated decrementation *terminates*. On the other hand, U consists of *infinite* or *non-terminating* numbers, so it is the canonical example of a non-overt subspace in recursion theory. Since U is not overt, it's not open, so it's not defined by a predicate. Hence D is not closed or compact, so we do not expect [J] to provide its supremum.

We do not know whether there is in fact a Dedekind-complete but non-Archimedean "real line" in ASD. This is a difficult but intriguing problem in recursion theory. Careful study of John Conway's construction may yield a recursive analogue (when this conjecture was put to him, he considered it plausible). The principal difficulty arises from the alternating quantifiers in the definition of $<$, as the arithmetic operations are clearly constructive [Ros01]. Even if such a model does exist, Stolz's argument would still show that the sequence 0, 1, ..., has no limit: any infinite element ω is "inaccessible" from finite values.

Such an object could, of course, be rather useful to develop differential calculus in a "non-standard" way [Koc06, Rob66]. It would also illustrate the importance of overtness very clearly. Here we have simply "left the door open" to such a possibility, by using approximate division and arithmetic locatedness in the proof, instead of the Archimedean principle itself.

Theorem 12.14 Let Q be any linearly ordered field (or commutative ring with approximate division) for which every Dedekind cut is arithmetically located. Then these cuts form a commutative ring. $\square$

13 Reciprocals and roots

Arithmetic, of course, provided one of the principal reasons for introducing Dedekind cuts in the first place, namely the solution of (algebraic and other) equations. Equations involving general continuous functions may be solved using the Intermediate Value Theorem, which we prove for ASD in [J].

In this section we present a simpler technique that manipulates Dedekind cuts directly. This is sufficient to define the inverse of a strictly monotone function, for example to find cube roots. However, it is more natural to present the method symmetrically in the given function and its inverse, and then $x \mapsto x^{5/3}$, exp and log are equally simple examples. A particular problem of this kind is then formulated as a pair of binary *relations*. Anyone with a little knowledge of lattice or category theory will recognise this situation as an *adjunction*, or rather a pair of them.

Definition 13.1 A ***strictly monotone graph*** $Q_1 \leftrightarrow Q_2$ between dense linear orders without endpoints (Definition 6.1) is a pair $(\lessdot, \gtrdot)$ of binary relations that satisfy

$$\begin{array}{rcccl} \exists b.\, a <_1 b \lessdot x & \Leftrightarrow & a \lessdot x & \Leftrightarrow & \exists y.\, a \lessdot y <_2 x \\ \exists b.\, x \gtrdot b <_1 a & \Leftrightarrow & x \gtrdot a & \Leftrightarrow & \exists y.\, x <_2 y \gtrdot a \\ \exists x.\, a \lessdot x \gtrdot b & \Leftrightarrow & a <_1 b \qquad x <_2 y & \Leftrightarrow & \exists a.\, x \gtrdot a \lessdot y \end{array}$$

where $a, b : Q_1$ and $x, y : Q_2$. These objects could, for example, consist of just *positive* rationals, or carry the opposite of the usual order. Let R_1 and R_2 be their Dedekind completions. Semi-continuous functions may be encoded in a similar way, but using just one of the two relations.

Proposition 13.2 Any strictly monotone graph defines inverse functions $R_1 \cong R_2$ by

$$\begin{array}{rcl} f(\delta, \upsilon) & \equiv & (\lambda e.\, \exists d.\, e \lessdot d \wedge \delta d, \quad \lambda t.\, \exists u.\, u \gtrdot t \wedge \upsilon u) \\ g(\epsilon, \tau) & \equiv & (\lambda d.\, \exists e.\, d \gtrdot e \wedge \epsilon e, \quad \lambda u.\, \exists t.\, t \lessdot u \wedge \tau t), \end{array}$$

for which $a \lessdot x \iff a <_1 fx \iff ga <_2 x$ and $x \gtrdot a \iff fx <_1 a \iff x <_2 ga$.

Proof The first four axioms make $f(\delta, \upsilon)$ and $g(\epsilon, \tau)$ rounded if (δ, υ) and (ϵ, τ) are. The other two transfer boundedness, disjointness and locatedness, and also make the maps inverse. □

We can use this to define reciprocals and roots, but since their domains are restricted and reciprocals reverse the order, we must modify Q and R before using them in the Proposition.

Lemma 13.3 Let Q be a linearly ordered ring that has approximate division, and let $Q_1 \equiv Q_2 \subset Q$ be its open subspace of positive elements. Then

$$a <_1 b \;\equiv\; a < b, \qquad x <_2 y \;\equiv\; x > y, \qquad a \lessdot x \;\equiv\; ax < 1, \qquad x \gtrdot a \;\equiv\; 1 < ax$$

define a strictly monotone graph $Q_1 \leftrightarrow Q_2$.

Proof Approximate division provides the six conditions

$$\begin{array}{rcccl} \exists b.\, a < b \wedge bx < 1 & \Leftrightarrow & ax < 1 & \Leftrightarrow & \exists y.\, ay < 1 \wedge y > x \\ \exists b.\, 1 < bx \wedge b < a & \Leftrightarrow & 1 < ax & \Leftrightarrow & \exists y.\, x > y \wedge 1 < ay \\ \exists x.\, ax < 1 < bx & \Leftrightarrow & a < b \qquad x > y & \Leftrightarrow & \exists a.\, ay < 1 < ax, \end{array}$$

in particular $0 < a < b \Rightarrow \exists x.\, a < abx < b \Rightarrow \exists x.\, bx > 1 > ax.$ □

Theorem 13.4 R is an ordered field, in which x^{-1} is defined for $x \neq 0$ by

$$\begin{array}{rcl} (\delta, \upsilon)^{-1} & \equiv & \big(\lambda d.\, \exists u.\, \upsilon u \wedge ((du < 1 \wedge \delta 0) \vee (1 < du \wedge d < 0)), \\ & & \lambda u.\, \exists d.\, \delta d \wedge ((du < 1 \wedge u > 0) \vee (1 < du \wedge \upsilon 0))\big). \end{array}$$

In the strictly positive or negative cases this is respectively

$$\big(\lambda d.\ (d<0)\vee(\exists u.\ \upsilon u\wedge du<1),\quad \lambda u.\ (u>0)\wedge(\exists d.\ \delta d\wedge du<1)\big)$$

or

$$\big(\lambda d.\ (d<0)\wedge(\exists u.\ \upsilon u\wedge 1<du),\quad \lambda u.\ (u>0)\vee(\exists d.\ \delta d\wedge 1<du)\big).$$

Proof The previous two results define an involution on $\{x:\mathbb{R}\mid x>0\}$. If Q is a field then this agrees with the reciprocal on Q and with Moore's formulae (Definition 2.3) in the legitimate case $0<a\le b$, where

$$(\delta_a,\upsilon_b)^{-1}\ =\ (\delta_{1/b},\upsilon_{1/a})\quad\text{and}\quad(\delta_{-b},\upsilon_{-a})^{-1}\ =\ (\delta_{-1/a},\upsilon_{-1/b}).$$

Without assuming that Q is a field, $(\epsilon_q,\tau_q)\equiv(\delta_q,\upsilon_q)^{-1}$ for $q>0$ is given by $\epsilon_q d\equiv(dq<1)$ and $\tau_q u\equiv(1<qu)$. The arithmetical laws follow from Lemma 11.18. The negative case is similar, and we observe that the given formula combines the two cases. □

Remark 13.5 We do not need to consider 0 this time, but since

$$\exists du.\ (\delta,\upsilon)^{-1}(d,u)\implies\delta 0\vee\upsilon 0,$$

the value in any illegitimate case, including 0^{-1} and $(\delta_{-1},\upsilon_1)^{-1}$, is $(\bot,\bot)$. This denotes the interval $[-\infty,+\infty]$.

More generally, the reciprocal of a general (and possibly back-to-front) interval with endpoints is given by

$[d,u]^{-1}$	$d<0$	$d=0$	$d>0$
$u<0$	$[u^{-1},\,d^{-1}]$	$[u^{-1},-\infty]$	$[+\infty,-\infty]$
$u=0$	$[-\infty,+\infty]$	$[-\infty,+\infty]$	$[+\infty,\,d^{-1}]$
$u>0$	$[-\infty,+\infty]$	$[-\infty,+\infty]$	$[u^{-1},\,d^{-1}]$,

illustrating the one-sided nature of Scott continuity: over $d\le 0$ for the ascending real number d and over $u\ge 0$ for the descending one u.

Theorem 13.6 If Q is Archimedean then so is R.

Proof Given $x,y:R$ with $y>0$, put $z\equiv x/y$. Then $z-\frac{1}{2}<d<z<u<z+\frac{1}{2}$ for some $d,u:Q$ with $u-d<1$. Then $n-1<d<z<u<n+2$ by the Archimedean principle for Q, whence either $(n-1)y<x<(n+1)y$ or $ny<x<(n+2)y$. □

Lemma 13.7 Q has ***approximate roots*** in the sense that, for $d,u:Q$ and $1\le n:\mathbb{N}$,

$$\begin{array}{lll}(0<d<u) & \Rightarrow & \exists x{:}Q.\ d<x^{2n}<u\\ (d<u) & \Rightarrow & \exists x{:}Q.\ d<x^{2n+1}<u.\end{array}$$

Proof These may be found in an Archimedean ordered commutative ring with approximate division by an algorithm similar to Proposition 12.7, for which we cite Babylonian clay tablets as the original source.

Alternatively, and without relying on the Archimedean principle, we may use the constructive approximate intermediate value theorem, namely

$$fe<d<u<ft\implies\exists x{:}[e,t].\ d<fx<u.$$

The usual proof is that the non-empty open subspaces

$$D\equiv\{x\mid fx<u\}\qquad\text{and}\qquad U\equiv\{x\mid d<fx\}$$

cover the interval $[e,t]$, and so must intersect. This is one of the ways of saying that the real interval is ***connected***, which we prove for ASD in [J]. As the intersection is open, it must contain a rational.

The result is applicable to $fx \equiv x^{2n}$ when we put

$$e \equiv \begin{cases} d & \text{if } 0 < d < 1 \\ \frac{1}{2} & \text{if } \frac{1}{2} < d \end{cases} \qquad \text{and} \qquad t \equiv \begin{cases} u & \text{if } 1 < u \\ 2 & \text{if } u < 2 \end{cases}$$

and to $fx \equiv x^{2n+1}$ when we put

$$e \equiv \begin{cases} d & \text{if } d < -1 \\ -2 & \text{if } -2 < d \end{cases} \qquad \text{and} \qquad t \equiv \begin{cases} u & \text{if } 1 < u \\ 2 & \text{if } u < 2. \end{cases} \qquad \square$$

Proposition 13.8 R has $(2n+1)$st roots and $[0,\infty)$ has $2n$th roots, where

$$\begin{aligned} \sqrt[2n+1]{(\delta,\upsilon)} &\equiv \big(\lambda d.\, \delta(d^{2n+1}),\ \lambda u.\, \upsilon(u^{2n+1})\big) \\ +\sqrt[2n]{(\delta,\upsilon)} &\equiv \big(\lambda d.\, (d<0) \vee \delta(d^{2n}),\ \lambda u.\, (u>0) \wedge \upsilon(u^{2n})\big). \end{aligned}$$

In the illegitimate case, $\sqrt{-1} = 0$.

Proof The strictly monotone graph is given by the relations $(a < x^m)$ and $(x^m < a)$, where $m \equiv 2n$ or $2n+1$. If $a < x^m$ and $0 < x$ then, by m-fold application of roundedness of division,

$$\exists 0 < y_1, y_2, \ldots, y_m < x.\ \mathsf{max}(0,a) < y_1 y_2 \cdots y_m < \cdots < y_1 y_2 x^{m-2} < y_1 x^{m-1} < x^m,$$

so $a < y^m$, where $0 < y \equiv \mathsf{max}(y_1, \ldots, y_m) < x$. Similarly if $0 < x$ and $x^m < a$ then $\exists y > x.\, y^m < a$, whilst if $m \equiv 2n+1$ and $x < 0$ then we switch the signs. Approximate roots provide the fifth interpolation property, and also the case $x \equiv 0$, since Q has decidable equality. The other conditions are trivial. $\square$

Finally, we leave the interested reader to use this method to develop the exponential and logarithm functions $\mathbb{R} \rightleftarrows (0,\infty)$ with base b, using $0 < d < b^{k \cdot 2^{-n}} < u$.

14 Axiomatic completeness

Our construction of the real line in ASD is finished. At any rate, we have found an object that has all of the properties that we listed in Definition 1.1, so we shall now write $\mathbb{R}$ for it.

Are these axioms sufficient to characterise the reals? We shall show that any object R that satisfies Definition 1.1 is uniquely isomorphic to the one that we have constructed.

However, we shall state the assumption that R be Dedekind-complete explicitly when we need it, and leave it out otherwise. This is because we also intend to explore the analogy between Dedekind completeness and definition by description, which are both special cases of sobriety (Definition 4.22). All three of these ideas say how the manipulation of logical expressions has an impact back on numbers.

We claim that ASD (in particular, sobriety) generalises Dedekind's intuition — to the extent of making Dedekind completeness *redundant*. This means that, if you try to encapsulate the *Cauchy* reals in some construction within ASD, the result will actually be *Dedekind* complete. We find the limit of any Cauchy sequence as a Dedekind cut in [J].

We begin with the well known argument for the uniqueness of $\mathbb{R}$ based on its algebra, order, Dedekind completeness and the Archimedean principle.

Proposition 14.1 Let R be an ordered field. Then there are unique monos $\mathbb{N} \to R$, $\mathbb{Z} \to R$ and

$j : \mathbb{Q} \to R$ that preserve 0, 1, $+$, $\times$, $<$ and (where appropriate) $-$ and $\div$. In particular, R has characteristic 0. $\square$

Lemma 14.2 If $d < u$ in an Archimedean ordered field R then there is some $q : \mathbb{Q}$ with $d < q < u$, and we may choose it to be dyadic, *i.e.* of the form $q \equiv k \cdot 2^{-n}$ with $k, n : \mathbb{Z}$.

Proof First apply the Archimedean axiom to $q \equiv u - d > 0$ and $p \equiv 2$, giving $m, n : \mathbb{Z}$ with

$$(m-1)(u-d) \ < \ 2 \ < \ (m+1)(u-d) \ < \ 2^n(u-d).$$

Then use it again with $q' \equiv 2^{-n}$, giving $k : \mathbb{N}$ with

$$d < u - 2 \cdot 2^{-n} < k \cdot 2^{-n} < u < (k+2) \cdot 2^{-n}. \qquad \square$$

Proposition 14.3 Let R be an Archimedean ordered field. Then there is a mono $\ell : R \rightarrowtail \mathbb{R}$ that fixes $\mathbb{Q}$, and preserves and reflects $<$.

Proof For $x : R$, consider $\delta_x, \upsilon_x : \Sigma^{\mathbb{Q}}$, defined by $\delta_x d \equiv (d <_R x)$ and $\upsilon_x u \equiv (x <_R u)$. Since $\mathbb{Q} \to R$ preserves and reflects order, for $x : \mathbb{Q}$, these formulae agree with Lemma 6.12.

They are clearly lower, upper and disjoint, using the basic properties of the order on R. As this order is located in the sense that $(d <_R u) \Rightarrow (d <_R x) \vee (x <_R u)$, the pair (δ_x, υ_x) is located in the sense of Definition 6.8. Boundedness follows directly from the Archimedean principle, and roundedness by Lemma 14.2.

Hence (δ_x, υ_x) is a Dedekind cut of $\mathbb{Q}$, and we have defined a map $\ell : R \to \mathbb{R}$. Using Lemma 14.2 for R again, together with Definition 6.10 for $<_{\mathbb{R}}$,

$$x <_R y \iff \exists q{:}\mathbb{Q}.\ x <_R q <_R y \ \equiv \ \exists q{:}\mathbb{Q}.\ \delta_x q \wedge \upsilon_x q \ \equiv \ \ell x <_{\mathbb{R}} \ell y,$$

so the order relations $<_R$ and $<_{\mathbb{R}}$ agree, and (as they are total) the map ℓ is mono. $\square$

Theorem 14.4 If R is a Dedekind-complete Archimedean ordered field then the map ℓ in the previous result is an isomorphism. Definition 1.1 therefore characterises $\mathbb{R}$ uniquely up to unique isomorphism.

Proof Members of $\mathbb{R}$ are by definition Dedekind cuts of $\mathbb{Q}$, which extend to cuts of R as in Proposition 9.5. These define *members* of R since it is Dedekind-complete. Since all of the orders agree, this process is inverse to the construction of ℓ. $\square$

Throughout this paper we have emphasised that we regard *topology* as primary, but it was not mentioned in the previous result. So we shall now investigate Σ^R, deducing that $\ell : R \cong \mathbb{R}$ from the principles of ASD, *without* assuming Dedekind completeness. The arguments are essentially those that led up to the formulation of $\mathcal{E}$ in Section 2.

Proposition 14.5 For $x : R$ and $\phi : \Sigma^R$, $\quad \phi x \iff \exists du.\ (d < x < u) \wedge \forall y{:}[d,u].\ \phi y.$

Proof Whilst Corollary 10.8 has already shown that the Dedekind reals $\mathbb{R}$ (as we have constructed them) satisfy this property, we claim here that it also follows from Definition 1.1. This is proved in detail in [J]. The key point is that, since $[-1,+1] \subset R$ is compact, Hausdorff and totally ordered,

$$\phi x \implies \forall h{:}[-1,+1].\ \exists \delta > 0.\ \big(|h| > \delta \vee \phi(x+h)\big),$$

and then Scott continuity (Axiom 4.21) allows us to interchange the quantifiers. This uses neither Dedekind completeness nor the Archimedean principle. $\square$

Lemma 14.6 The quantifier $\exists_R$ is given by $\exists_{\mathbb{Q}}$, as in Theorem 9.2.

Proof By Proposition 14.5 and Lemma 14.2,

$$\exists x{:}R.\ \phi x \ \Rightarrow\ \exists dux{:}R.\ (d < x < u) \wedge \forall y{:}[d,u].\ \phi y \ \Rightarrow\ \exists q{:}Q.\ \phi(jq) \ \Rightarrow\ \exists x{:}R.\ \phi x. \qquad \square$$

Proposition 14.7 $\Sigma^\ell : \Sigma^{\mathbb{R}} \cong \Sigma^R$.

Proof As in Proposition 2.15, we define, for $\phi : \Sigma^R$,

$$I\phi \ \equiv\ \lambda\delta\upsilon{:}\Sigma^{\mathbb{Q}}.\ \exists d < u{:}\mathbb{Q}.\ \delta d \wedge \upsilon u \wedge \forall y{:}[d,u].\ \phi y : \Sigma^{\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}},$$

using the Heine–Borel property that we have assumed for $[d,u] \subset R$. Then

$$I\phi(\ell x) \ \equiv\ I\phi(\delta_x, \upsilon_x) \ \equiv\ \exists du.\ (d <_R x <_R u) \wedge \forall y{:}[d,u].\ \phi y \ \Leftrightarrow\ \phi x,$$

in which d and u may be chosen from either $\mathbb{Q}$ or R, by Lemma 14.6, and we recognise the formula in Proposition 14.5. So $\Sigma^\ell \cdot I$ is the identity on Σ^R.

Now, the proof of Proposition 2.17 (or its modification for Notation 8.2) is also valid for R in ASD. Hence Σ^R and $\Sigma^{\mathbb{R}}$ split the same idempotent $\mathcal{E}$ on $\Sigma^{\Sigma^{\mathbb{Q}}\times\Sigma^{\mathbb{Q}}}$, *cf.* Remark 5.2, so they are isomorphic. $\square$

The topology on $\mathbb{R}$ is therefore unique, when we assume the Heine–Borel property but not necessarily Dedekind completeness. As in the title of Dedekind's paper, the topology comes first, and we derive irrational numbers from that.

Theorem 14.8 Consider Definition 1.1, including the Heine–Borel property but omitting Dedekind completeness. If R satisfies this and is also sober, or indeed definable, then it is uniquely isomorphic to $\mathbb{R}$.

Proof Propositions 4.23, 4.25, 5.18 and 14.7. $\square$

Remark 14.9 On the other hand, we could bypass the constructions of this paper by defining $\mathbb{R}$ as a *base* type, together with (all of) the structure of Definition 1.1.

In Axiom 4.24 we showed how descriptions for $\mathbb{N}$ may be formulated in a type-theoretic style by giving the introduction rule for "the". In this, the elimination, β- and η-rules are given by singleton and substitution, and there is a normalisation theorem [A, §8]. It is easy to overlook the key role of equality ($=$) in this — it is to descriptions as application is to λ-abstraction and membership is to set-formation. Since $\mathbb{R}$ is Hausdorff and not discrete, we cannot base an analogous treatment of it on equality. We could use *in*equality ($\neq$) (Exercise 14.14), but it is more natural to consider $<$ and $>$, as these give rise to the lower and upper parts of a Dedekind cut.

Definition 14.10 Dedekind cuts and the arithmetical order may be seen as introduction and elimination rules for the type $\mathbb{R}$.

(a) $\mathbb{R}$-***introduction***: given a cut $(\delta,\upsilon) : \Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}$, we introduce a real number $\mathsf{cut}(\delta,\upsilon) : \mathbb{R}$;

(b) $\mathbb{R}$-***elimination***: given a real number $a : \mathbb{R}$ and rationals $d, u : \mathbb{Q}$, we eliminate a in favour of $(d < a), (a < u) : \Sigma$, along with the other rules concerning the order $<$;

(c) $\mathbb{R}$-***beta***: $\delta d \Leftrightarrow \big(d < \mathsf{cut}(\delta,\upsilon)\big)$ and $\upsilon u \Leftrightarrow \big(\mathsf{cut}(\delta,\upsilon) < u\big)$;

(d) $\mathbb{R}$-***eta***: $a = \mathsf{cut}(\lambda d.\ d < a, \lambda u.\ a < u)$.

Whenever we add new features like this, we have to ask what effect they have on the terms and equations of the existing calculus.

Theorem 14.11 $\mathbb{R}$ is a conservative extension of the ASD calculus.

Proof As we have already constructed an object that satisfies Definitions 1.1 and 14.10 within the calculus, the earlier results of this section show that the new object is uniquely isomorphic to the old one. Similarly, any new type expressions are isomorphic to old ones. These isomorphisms translate the new terms of the old types into provably equal old terms, and the new equations are also redundant. This means that the new category is *equivalent* to the old one. □

Remark 14.12 As this argument is not very illuminating, here is an explicit procedure for eliminating (Dedekind) cut operations, except on the outside of a term of type $\mathbb{R}$.
(a) If we are given a term a of type $\mathbb{R}$, we use the $\mathbb{R}$-η-law to wrap it in a cut:

$$a \;=\; \mathsf{cut}\,(\lambda d.\, d < a,\ \lambda u.\, a < u).$$

(b) Any cut that occurs as a proper sub-term of this is itself of type $\mathbb{R}$, but is enclosed in a smallest sub-term σ of type Σ. By Proposition 14.5, this is of the form

$$[\mathsf{cut}(\delta, \upsilon)/x]^*\sigma \iff (\lambda x.\, \sigma)\,\mathsf{cut}(\delta, \upsilon) \iff \exists d < u{:}\mathbb{Q}.\ \delta d \wedge \upsilon u \wedge \forall x{:}[d, u].\ \sigma.$$

This normalisation theorem means that we compute with logical expressions (involving predicates on rationals) instead of real numbers. Notice that Dedekind completeness and the Heine–Borel theorem are intimately related. In fact, instead of the quantifier $\forall x : [d, u]$, we may simply evaluate σ according to Moore's rules at the interval $\mathsf{x} \equiv [d, u]$. This is part of a translation that has the effect of replacing continuous variables by intervals, which is what programmers do anyway, but it is justified formally in [K]. □

As $\mathbb{R}$ is no longer one of the types that are generated from $\mathbb{N}$ using powers and Σ-split subspaces, we need to show that it is sober, *cf.* Proposition 4.25(b) for $\mathbb{N}$.

Lemma 14.13 If $P : \Sigma^{\Sigma^R}$ is prime then

$$(\delta, \upsilon) \;\equiv\; \big(\lambda d.\, P(\lambda x.\, d < x),\ \lambda u.\, P(\lambda x.\, x < u)\big)$$

is a Dedekind cut, with $\mathsf{cut}(\delta, \upsilon) = \mathsf{focus}\, P$, so

$$(d < \mathsf{focus}\, P) \;\Leftrightarrow\; P(\lambda x.\, d < x) \quad \text{and} \quad (\mathsf{focus}\, P < u) \;\Leftrightarrow\; P(\lambda x.\, x < u).$$

Proof The last part is equivalent to the definition of δ and υ. These are rounded and bounded because

$$P(\lambda x.\, d < x) \;\Leftrightarrow\; P(\lambda x.\, \exists e.\, d < e < x) \;\Leftrightarrow\; \exists e.\, d < e \wedge P(\lambda x.\, e < x)$$

and

$$\top \;\Leftrightarrow\; P\top \;\Leftrightarrow\; P(\lambda x.\, \exists e.\, e < x) \;\Leftrightarrow\; \exists e.\, P(\lambda x.\, e < x),$$

using inter- and extrapolation and Scott continuity. By Lemma 4.9, with $\sigma \equiv (d < u)$,

$$P(\lambda x.\, d < u) \iff P\bot \vee (d < u) \wedge P\top \iff (d < u),$$

since P preserves $\top$ and $\bot$. Then (δ, υ) are disjoint because P preserves $\wedge$,

$$P(\lambda x.\, d < x) \wedge P(\lambda x.\, x < u) \;\Leftrightarrow\; P(\lambda x.\, d < x < u) \;\Rightarrow\; P(\lambda x.\, d < u) \;\Leftrightarrow\; (d < u),$$

and, by the dual argument, located because P preserves $\vee$:

$$P(\lambda x.\, d < x) \vee P(\lambda x.\, x < u) \;\Leftrightarrow\; P(\lambda x.\, d < x \vee x < u) \;\Leftarrow\; P(\lambda x.\, d < u) \;\Leftrightarrow\; (d < u). \qquad \square$$

Exercise 14.14 Using $\neq$ instead, we can develop ***definition by co-description*** for real numbers, replacing the singleton with its complement. But since $\exists$ also becomes $\forall$, we have to consider this in a compact Hausdorff space such as $\mathbb{I}$, instead of $\mathbb{R}$. The existence and uniqueness conditions for ω to classify a co-singleton are then the duals of those in Axiom 4.24:

$$\bot \iff \forall x{:}[0,1].\ \omega x \qquad \text{and} \qquad x \neq y \Longrightarrow \omega x \vee \omega y.$$

Consider any function $f : [0,1] \to \mathbb{R}$. Use the Heine–Borel Theorem to show that $\omega x \equiv \exists z.\ fz < fx$ satisfies the first condition, and hence that $\delta d \equiv \forall x.\ (d < x) \vee \omega x$ and $\upsilon u \equiv \forall x.\ (x < u) \vee \omega x$ define an interval, *i.e.* they are rounded, bounded and disjoint. Classically, this interval is the convex hull of the set on which f attains its minimum value.

Now suppose that $(x \neq y) \Longrightarrow \exists z.\ (fz < fx) \vee (fz < fy)$, *cf.* [Sch03]. This is the uniqueness condition for ω. Use Lemma 4.13 for compact Hausdorff spaces to deduce that (δ, υ) is located, giving a Dedekind cut $w \in [0,1]$. Show that w is the minimum of f, *i.e.* it satisfies $(x \neq w) \Rightarrow (fw < fx)$.

Show that definition by co-description is equivalent to sobriety and Dedekind completeness, using the translations

$$\omega \ \equiv\ \delta \vee \upsilon \qquad \text{and} \qquad P\phi \ \equiv\ \exists d < u.\ \delta d \wedge \upsilon u \wedge \forall x{:}[d,u].\ \phi x.$$

Expressing any co-description ω as a Dedekind cut is a special case of the characterisation of connected subspaces of $\mathbb{R}$ in [J].

From roundedness, boundedness and disjointness, we easily deduce that P preserves $\bot$, $\top$ and $\wedge$; these also entail the existence condition for the co-description ω, but this is more difficult. On the other hand, locatedness easily gives the uniqueness condition for ω, whilst it is more difficult to show that P preserves $\vee$. $\square$

15 Recursive analysis

Whether the closed interval is, or ought to be, compact in computable and constructive mathematics is a question that never ceases to be interesting. This is partly because compactness is such an important property for a topological space to have, and partly because there are several schools of constructive and computable mathematics that disagree about the answer, each of which sets up its own framework, differing slightly from all the others. The matter is made all the more confusing by the number of different definitions of compactness that have been given, even classically [Bou66, §I 9]. In this section we make a few observations about compactness of the closed interval in ASD and how it relates to various other schools of computability.

Most urgently, we hear alarm bells ringing in the minds of recursive analysts. On the one hand, we saw in Remark 3.4 that there is a model of ASD in which maps are precisely the computable maps between computably based locally compact spaces. On the other, Section 10 proved that the closed interval is compact. If all functions are computable, then surely we should be able to use some enumeration of the definable points of the closed interval to show that it is *not* compact?

Remark 15.1 In Recursive Analysis, the closed interval $[0,1]$ can be shown *not* to be compact by means of a ***singular cover***, *i.e.* a countable cover by open intervals with rational endpoints whose total length is bounded above by $\frac{1}{2}$. It can be shown that no finite sub-cover of a singular cover is a cover of $[0,1]$, see [TvD88, §6.4.3] or [BR87, §3.4]. Such covers exist in Recursive Analysis [TvD88, §6.4.2] because of the *formal Church's Thesis* CT_0 [TvD88, §3.3.2], which states that every function $\mathbb{N} \to \mathbb{N}$ has a Gödel code. (With hindsight, it would have been more natural to enumerate $\mathbb{N} \to \mathbb{N}_\bot$, but this is the way in which CT_0 is usually stated.) It follows that every real number has a Gödel code, too. In classical recursion theory, a singular cover is described as a computable enumeration of open intervals with rational endpoints of bounded total length which effectively covers all computable reals.

Example 15.2 The closed interval may also fail to be compact for reasons that have nothing to do with computability. The construction of the "Dedekind reals" R as the equaliser in Remark 6.9 can be done in any category with finite limits, certain exponentials and some mild assumptions on $\Sigma^{\mathbb{N}}$, as in Section 4. An example is the category **Dcpo** of posets with directed joins and Scott-continuous functions. Here the equaliser carries the discrete ("specialisation") order, but in this category the order determines the topology. The object R therefore also carries the discrete topology, so Σ^R is the powerset $\mathcal{P}(R)$ with the Scott topology, and the only compact subsets of R are the finite ones. Hence the predicate $[d,u] \subset U$ is not continuous in U. The formula in Proposition 2.15 gives an element of $\Sigma^{\Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}}$ for each $U \subset R$, but the function I is not Scott-continuous, so it does not yield a sound interpretation of the ASD calculus (Warning 5.17).

Remark 15.3 One resolution of the abnormalities of Recursive Analysis that has been adopted by several schools is that of *relative computability*, in which the *maps* are computable, but the *spaces* are the classical topological ones, *i.e.* they have "all" points instead of just the computable ones. Each space is *also* equipped with a "computability" structure, such as an enumeration of basic open sets, that is used to specify which continuous maps are computable. Examples are (one variant of) Type Two Effectivity [Wei00], and computable equilogical spaces [Bau00]. Another is Martín Escardó's synthetic topology of data types [Esc04], but beware that, whilst this is very similar to the calculus in Section 4, it lacks the crucial monadic property of ASD (Section 5). In all of these three settings, the closed interval is compact, the proof being the well known classical one. In particular, Escardó stresses the need for "all" points, as compactness would fail in a purely recursive version of his theory.

Remark 15.4 The sceptical pupil from the Russian School will not be so easily convinced, but will press us on the structure of the *syntactic* term model of ASD, in which the objects are types and the morphisms are the definable maps. Since everything is enumerable in such a model, one might expect to obtain a singular cover, and so non-compactness of the closed interval. Indeed, by following the usual construction, we could define a sequence of intervals with rational endpoints, (a_n, b_n), whose total length is bounded by $\frac{1}{2}$, and prove the meta-theorem that

$$\text{“if} \quad \vdash t : [0,1] \quad \text{then} \quad \vdash \exists n{:}\mathbb{N}.\ a_n < t < b_n\text{”}.$$

However, it does not follow that (a_n, b_n) *covers* $[0,1]$, by which we mean

$$x : [0,1] \ \vdash\ \exists n{:}\mathbb{N}.\ a_n < x < b_n,$$

because that would be to confuse a *family* of theorems about *all definable closed terms* t with a *single* theorem containing a free variable x. In category theory, the calculus is said to be *not well pointed*. Even though its morphisms are recursively enumerable, there are not enough of them $\mathbf{1} \to R$ to cover the equaliser.

Example 15.5 As a final attempt to break compactness of the closed interval in ASD, we might try to interpret it in a setting, such as Markov's Recursive Mathematics, in which the formal Church's Thesis CT_0 holds. A useful formulation of this is the category **PER** of partial equivalence relations over Stephen Kleene's first algebra [Kle45], or Martin Hyland's larger *effective topos*, **Eff** [Hyl82].

We try to interpret ASD in **PER** or **Eff** in the natural way, using the universal properties of $\mathbb{N}$, products, equalisers and exponentials. For Σ we take the object of semidecidable propositions.

This yields a valid interpretation of the language and construction of Sections 4 and 6, so that the equaliser R in Remark 6.9 is perhaps a reasonable candidate for the "real line" in these worlds.

However, we do not have a sound interpretation of Σ-split subspaces, because the map I that Remark 5.17 requires need not exist. In particular, the object Σ^R does not have the properties of the "Euclidean" topology, and the closed interval is not compact.

In **PER**, the formula in Proposition 2.15 does not define a morphism I because $[d, u] \subset U$ is not a completely r.e. predicate (in d, u and U). In the internal language of **Eff**, the sub-expression $\forall x{:}[d, u].\ x \in U$ is a term of type Ω but not Σ. Consequently, the formula defines a morphism I from Σ^R to $\Omega^{\Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}}$ and not to $\Sigma^{\Sigma^{\mathbb{Q}} \times \Sigma^{\mathbb{Q}}}$.

This situation is better documented for Cantor space:

Example 15.6 Richard Friedberg [Fri58] defined an effective operation on the set of total recursive functions $\mathbb{N} \to \mathbb{N}$ that is not the restriction of a recursive functional. See [Rog92, §15.3.XXXI] for another account of this construction, which also works for the total recursive binary sequences $\mathbb{N} \to \mathbf{2}$.

In the effective topos we can reformulate Friedberg's example by saying that the topology of $\mathbf{2}^{\mathbb{N}}$ (*i.e.* the object $\Sigma^{\mathbf{2}^{\mathbb{N}}}$) is not the subspace topology induced by the inclusion $\mathbf{2}^{\mathbb{N}} \subset \mathbf{2}_{\perp}^{\mathbb{N}}$. In this topos, $\mathbf{2}^{\mathbb{N}}$ is actually homeomorphic to $\mathbb{N}^{\mathbb{N}}$, so neither of these spaces can be locally compact, and their topology is not the subspace topology of $\mathbf{2}_{\perp}^{\mathbb{N}}$.

In the construction of $\mathbf{2}^{\mathbb{N}}$ in ASD [L], on the other hand, the space $\mathbf{2}_{\perp}^{\mathbb{N}}$ of partial functions from $\mathbb{N}$ to $\mathbf{2}$ plays a role similar to that of IR in Sections 7–8, being a closed subspace of $\Sigma^{\mathbb{N}} \times \Sigma^{\mathbb{N}}$. Then $\mathbf{2}^{\mathbb{N}} \subset \mathbf{2}_{\perp}^{\mathbb{N}}$ is defined by a nucleus that satisfies both the ASD and localic definitions. The quantifier $\forall$ that says that $\mathbf{2}^{\mathbb{N}}$ is compact is obtained from this nucleus in a similar way to that for $[d, u]$.

Escardó [Esc04] has shown how $\forall$ for $\mathbf{2}^{\mathbb{N}}$ may be implemented as a program in HASKELL. He proved its correctness using relative computability (Remark 15.3), but this may also be done using the monadic principle in ASD.

Remark 15.7 You may be left thinking that models of ASD are complicated, mysterious and hard to find. In fact, the calculus has interpretations in many kinds of classical or constructive set, type or topos theory. It is more accurate to say that these interpretations can be *based on* other foundational systems, because we obtain a model of our calculus by doing a *construction* on top of a system with weaker properties. In other words, whilst the direct interpretation in **PER** and **Eff** in Example 15.5 does not work, a slightly more involved one does.

Theorem 15.8 Let $(\mathcal{C}, \Sigma)$ be a model of the axioms in Section 4, in particular with products and powers Σ^X. Suppose also that idempotents split in $\mathcal{C}$. Then $\mathcal{A}^{\mathsf{op}}$ is a model of ASD, where $\mathcal{A}$ is the category of Eilenberg–Moore algebras for the monad arising from $\Sigma^{(-)} \dashv \Sigma^{(-)}$ [B]. □

Remark 15.9 Any category that results from this construction contains (the sober objects of) the original one (and more), and the embedding preserves $\Sigma^{(-)}$. However, many other constructions are *not* preserved, important examples being, of course, Cantor space and the Dedekind reals as we have defined them.

Whilst these *objects* may have had undesirable properties in the original structure, their *analogues* in the new one turn out to behave as the mainstream mathematician would expect. The new objects are by definition the exponentials, equalisers, *etc.* for the same data in the *new* category, whilst (the images of) the old objects are relieved of their former duties.

This construction can be carried out for all of the well known models of computable mathematics, including Recursive Mathematics, domain-theoretic **PER** models and Type Two Effectivity. For example, when we apply it to the category **CL** of continuous lattices and Scott-continuous functions, we obtain the category **LKLoc** of locally compact locales. It would be interesting and fruitful to do this in detail for the other cases, and, in particular, to compare their collections of overt objects.

On the other hand, these models have other significant properties besides those given in Section 4, which it may not be easy to reproduce in the new category. For example, **CL** is cartesian closed, but **LKLoc** is not.

Theorem 15.10 Over the effective topos **Eff** there is a sheaf topos $\mathbf{Eff}^{\mathsf{A}}$ that has a model $\mathcal{A}^{\mathsf{op}}$ of ASD as a reflective subcategory.[4]

Proof There is an *internal* category $\mathsf{C} \in \mathbf{Eff}$ of PERs that is weakly equivalent to a *reflective* subcategory $\mathcal{C} \subset \mathbf{Eff}$ of $\Omega_{\neg\neg}$-discrete objects [HR90]. (This situation is a peculiarity of realisability toposes, and the weakness of the equivalence is essential: Peter Freyd observed in the 1960s that, classically, any small complete category is a lattice [Tay99, Example 7.3.2(k)], and this extends to Grothendieck toposes.)

Now let A be the internal Eilenberg–Moore category for the monad over C within **Eff**, and $\mathcal{A}$ the external one over $\mathcal{C}$ (or equivalently over **Eff**). Then by Theorem 15.8, both A^{op} and $\mathcal{A}^{\mathsf{op}}$ are models of ASD.

Since A^{op} is an internal category, *i.e.* a *small* one in classical language, we may form the sheaf topos $\mathbf{Eff}^{\mathsf{A}}$. But A^{op} and $\mathcal{A}^{\mathsf{op}}$ are weakly equivalent to each other, and therefore to a full subcategory of $\mathbf{Eff}^{\mathsf{A}}$, by the Yoneda embedding. On the other hand, $\mathcal{A}^{\mathsf{op}}$ is also *totally cocomplete* (see [Kel86, Corollary 6.5] in particular, but the theory was developed in [SW78, Str80, Tho80, Woo82]), so this full subcategory is *reflective* in $\mathbf{Eff}^{\mathsf{A}}$. $\square$

Remark 15.11 Altogether, it is often very interesting (and not just for gladiatorial reasons) to investigate the conflicts between one view of the mathematical world and another, because the resolution of such conflicts often leads to powerful applications. It is therefore perhaps a little disappointing that we have not managed to bring ASD into confrontation, let alone conflict, with singular covers and similar recursive phenomena.

The explanation is simply that, as we have shown, the quantifier $\forall$ that encodes compactness of the closed interval is inter-definable with the map I that we introduced classically in Proposition 2.15. The systems in which the Heine–Borel theorem is true are exactly those in which I is definable. In any particular system, there may be some difficult proof of the existence of $\forall$ and I, but in ASD, the latter is given, essentially axiomatically, by the monadic principle.

16 Conclusion

We showed in Section 14 that Definition 1.1 is ***complete*** in the sense that it characterises $\mathbb{R}$ uniquely up to unique isomorphism of its structure.

Beware that, just as in the rest of the paper, this is a theorem *within* the calculus, about *each model* of ASD. In particular, our axioms isolate $\mathbb{R}$ within the classical model (Definition 3.1), where, of course, we intend the arithmetical operations and all other terms to be continuous. Maybe you believe that the real line satisfies some logically more powerful properties, such as excluded middle, the axiom of choice, the axiom-scheme of replacement, large cardinals, *etc.* These would justify the *existence* of many more points, functions and subspaces, and also prove more *equations* amongst them. Nevertheless, the uniqueness result still stands within your particular model of topology. The additional properties are axioms that we might add to ASD's account of topology as a whole, not to that of the real line within it.

[4]This is unpublished joint work that Giuseppe Rosolini and Paul Taylor did in 1997. However, current work by Rosolini that uses groupoids instead of PERs may improve on these results.

Questions of what is "true" in the Platonist mathematical world are unanswerable, but we can say something more definite for computation: ASD is itself a complete axiomatisation of countably based locally compact spaces [G].

Suppose that you have a continuous function $f : \mathbb{R} \to \mathbb{R}$ that is defined within your Platonist world, together with a program π taking three rational arguments such that $\pi(q, d, u)$ terminates iff $d < f(q) < u$. Then π may be translated into a term of type $\mathbb{R} \to \mathbb{R}$ in ASD, whose classical interpretation is the function f. Conversely, any term of ASD is a program (Remark 3.4).

It is therefore reasonable to ask whether our axioms also tell the whole story as far as analysis is concerned. We are making the bold (essentially philosophical) claim that the exceedingly weak computational logic of ASD is enough for "what matters". The answer must be a utilitarian one, in which we accumulate evidence that our object and the category in which it lives have many of the familiar properties of analysis.

The paper that follows [J] begins to answer this by studying some of the basic principles of analysis on the real line, such as convergence of Cauchy sequences, maxima of compact overt subspaces and connectedness, culminating in the intermediate value theorem. With the benefit of the Heine–Borel definition of compactness and the dual notion of overtness, we can develop these ideas in a *topological* style, in contrast to the *metrical* one that Bishop and others used.

Concentrating on general topology, one experimental test of whether we have "the real" real line is whether open subsets look like what we expect. Traditionally, any open subset of the real line is uniquely expressible as a countable union of disjoint open intervals. After some constructive re-interpretation of this statement, we can indeed prove it in ASD, but it depends on compactness of $[0, 1]$.

Our claim should be seen alongside the analogous one of Errett Bishop [Bis67]. He was not only constructive but also *conservative*, in the sense that his theorems are compatible with classical analysis (using a kind of constructive set theory [Bri99]), as well as with Brouwer's Intuitionism and Markov's Recursive Mathematics. Since these settings disagree about compactness of the closed interval, Bishop *omitted* it and focused on Cauchy completeness and total boundedness instead. The decomposition of open subsets of $\mathbb{R}$ into disjoint open intervals also fails.

Given that ASD *includes* the Heine–Borel theorem, Bishop's followers may jump to the conclusion that our theory is unacceptable in their system. Before they do so, they should first remember that the spaces in ASD are *not* sets and that the ASD calculus is *not* the calculus of logical predicates about points. The ASD statement "$[0, 1]$ is compact" says nothing about the closed interval $[0, 1]$ from Bishop's mathematics, because the ASD type R cannot be interpreted naïvely as Bishop's set of real numbers.

One way of understanding the logical strength of ASD is as an algebraic formulation of topology (Theorem 15.8). When the "ASD algebras" are constructed in Bishop's world, one of them will be called $[0, 1]$, and there will be an "ASD homomorphism" $\forall_{[0,1]} : \Sigma^{[0,1]} \to \Sigma$ witnessing the fact that $[0, 1]$ behaves like a compact space within the realm of ASD algebras. The interesting question from the point of view of a Bishop-style constructivist is not whether ASD algebras are acceptable *per se* (which they are, so far as we can tell), but how good a formulation of topology ASD provides.

On the other hand, ASD is not presented as algebra but as topology. This has been done by taking advantage of forty years' study of categorical logic, type theory and the relationship between the two. Using this, a terse piece of category theory has been transformed into a symbolic calculus that can be used by real analysts. If, despite our advocacy of the intrinsic nature of the topology of the real line, you still believe that sets of points are fundamental, then you will inevitably regard these "algebraic spaces" (along with locales and formal topology) as artificial. People once said the same thing about complex numbers.

However, this point of view is unsustainable in computation, where *everything* is a mirage — our calculus has at least as good a claim to represent real analysis as does one that is based on Gödel numberings of points, especially as we have presented it entirely syntactically.

Another point that might interest Bishop-style constructivists is our basic type Σ. Bishop — and he is certainly not the only one to do so — took *numbers* as fundamental, and proceeded to construct everything else from them. This is supposedly because integers are the basic objects of computation.

But modern computer science (in particular domain theory) teaches us that *observable properties* are equally, if not more, fundamental. The space Σ *is* the space of observable propositions. How might Bishop have axiomatised Σ, and, more generally, spaces of observable properties Σ^X?

Although we ourselves were brought up with domain-theoretic ideas, we had no intention of repeating them unnecessarily when we began our investigation of analysis. We just wanted to prove some very traditional results (Heine–Borel, boundedness of a function on a closed interval, and the intermediate value theorem), in order to show that ASD is capable of doing this.

However, we found that it was essential to define the ascending and descending reals and the interval domain *first*. Maybe you are interested in some other system (even the classical one) and have not followed any of the details of the construction in ASD. Nevertheless, a conclusion that we hope that you will take away from this paper is that $\underline{\mathbb{R}}$, $\overline{\mathbb{R}}$ and $\mathrm{I}\mathbb{R}$ are *more fundamental* than the familiar ("Euclidean") real line, *and not derived from it.*

Interval computation, and to a lesser extent the interval domain, are now common currency amongst those who program with real numbers in a way that is provably correct. As these people already know very well, computing a function *à la* Moore with an interval $[d, u]$ *over*estimates the image of the function on this interval. In our notation, $\Phi[d, u] \Rightarrow \forall x\!:\![d, u].\ \Phi(ix)$.

Dually, computing with the back-to-front interval $[u, d]$ *under*estimates the image, or as we would say, $\exists x\!:\![d, u].\ \Phi(ix) \Rightarrow \Phi[u, d]$. This is not so well known amongst interval analysts, or very well explained where it is known. It is clearly a conceptual error to regard the interval $[d, u]$ as the set $\{x \in \mathbb{R} \mid d \le x \le u\}$ — both forward and reverse intervals should be seen as generalised Dedekind cuts in the way that we described in Section 2.

Finally, we invite you to give careful consideration to our Definition 1.1 as the syntax of a language for analysis. We feel that this is worthy of study, both as the basis of theory and as that of computation, and hope that it will ultimately unite these disciplines in a constructive way.

References

[Bau00] Andrej Bauer. *The Realizability Approach to Computable Analysis and Topology.* PhD thesis, Carnegie Mellon University, 2000. Technical report CMU-CS-00-164.

[Bau08] Andrej Bauer. Efficient computation with Dedekind reals. In Vasco Brattka, Ruth Dillhage, Tanja Grubba, and Angela Klutch, editors, *Fifth International Conference on Computability and Complexity in Analysis*, Hagen, Germany, August 2008.

[BB85] Errett Bishop and Douglas Bridges. *Constructive Analysis.* Number 279 in Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1985.

[BB08a] Josef Berger and Douglas Bridges. The anti-Specker property, a Heine–Borel property, and uniform continuity. *Archive for Mathematical Logic*, 46:583–592, 2008.

[BB08b] Josef Berger and Douglas Bridges. The fan theorem and positive-valued uniformly continuous functions on compact intervals. *New Zealand Journal of Mathematics*, 38:129–135, 2008.

[Bis67] Errett Bishop. *Foundations of Constructive Analysis.* Higher Mathematics. McGraw–Hill, 1967.

[BK08] Andrej Bauer and Iztok Kavkler. Implementing real numbers with RZ. *Electronic Notes in Theoretical Computer Science*, 202:365–384, 2008.

[Bou66] Nicolas Bourbaki. *Topologie Générale*. Hermann, 1966. English translation, *General Topology*, distributed by Springer-Verlag, 1989.

[BR87] Douglas Bridges and Fred Richman. *Varieties of Constructive Mathematics*. Number 97 in London Mathematical Society Lecture Notes. Cambridge University Press, 1987.

[Bri99] Douglas Bridges. Constructive mathematics: a foundation for computable analysis. *Theoretical Computer Science*, 219:95–109, 1999.

[BS08] Andrej Bauer and Christopher Stone. RZ: a tool for bringing constructive and computable mathematics closer to programming practice. *Journal of Logic and Computation*, 19:17–43, 2008.

[Cle87] John Cleary. Logical arithmetic. *Future Computing Systems*, 2(2):125–149, 1987.

[CN96] Jan Cederquist and Sara Negri. A constructive proof of the Heine–Borel covering theorem for formal reals. In Stefano Beradi and Mario Coppo, editors, *Types for Proofs and Programs*, number 1158 in Lecture Notes in Computer Science. Springer-Verlag, 1996.

[Con76] John Horton Conway. *On Numbers and Games*. Number 6 in London Mathematical Society Monographs. Academic Press, 1976. Revised edition, 2001, published by A K Peters, Ltd.

[Ded72] Richard Dedekind. *Stetigkeit und irrationale Zahlen*. Braunschweig, 1872. Reprinted in [Ded32], pages 315–334; English translation, *Continuity and Irrational Numbers*, in [Ded01].

[Ded01] Richard Dedekind. *Essays on the theory of numbers*. Open Court, 1901. English translations by Wooster Woodruff Beman; republished by Dover, 1963.

[Ded32] Richard Dedekind. *Gesammelte mathematische Werke*, volume 3. Vieweg, Braunschweig, 1932. Edited by Robert Fricke, Emmy Noether and Øystein Ore; republished by Chelsea, New York, 1969.

[ES98] Abbas Edalat and Philipp Sünderhauf. A domain-theoretic approach to real number computation. *Theoretical Computer Science*, 210:73–98, 1998.

[Esc04] Martín Escardó. Synthetic topology of data types and classical spaces. *Electronic Notes in Theoretical Computer Science*, 87:21–156, 2004.

[FH79] Michael Fourman and Martin Hyland. Sheaf models for analysis. In Michael Fourman, Chris Mulvey, and Dana Scott, editors, *Applications of Sheaves*, number 753 in Lecture Notes in Mathematics. Springer-Verlag, 1979.

[Fri58] Richard Friedberg. Un contre-example relatif aux fonctionnelles récursives. *Comptes rendus hebdomadaires des scéances de l'Académie des Sciences (Paris)*, 247:852–854, 1958.

[Gen35] Gerhard Gentzen. Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1935. English translation in [Gen69], pages 68–131.

[Gen69] Gerhard Gentzen. *The Collected Papers of Gerhard Gentzen*. Studies in Logic and the Foundations of Mathematics. North-Holland, 1969. Edited by Szabo, Manfred.

[GHK+80] Gerhard Gierz, Karl Heinrich Hoffmann, Klaus Keimel, Jimmie Lawson, Michael Mislove, and Dana Scott. *A Compendium of Continuous Lattices*. Springer-Verlag, 1980. Second edition, *Continuous Lattices and Domains*, published by Cambridge University Press, 2003.

[Hey56] Arend Heyting. *Intuitionism, an Introduction*. Studies in Logic and the Foundations of Mathematics. North-Holland, 1956. Third edition, 1971.

[HLPZ] Guillaume Hanrot, Vincent Lefèvre, Patrick Pélissier, and Paul Zimmermann. *The MPFR Library*. INRIA. `www.mpfr.org`.

[HR90] Martin Hyland and Giuseppe Rosolini. The discrete objects in the effective topos. *Proceedings of the London Mathematical Society*, 60:1–36, 1990.

[Hyl82] Martin Hyland. The effective topos. In Anne Troelstra and Dirk van Dalen, editors, *The L.E.J. Brouwer Centenary Symposium*. North Holland, 1982.

[Hyl91] Martin Hyland. First steps in synthetic domain theory. In Aurelio Carboni, Maria-Cristina Pedicchio, and Giuseppe Rosolini, editors, *Category Theory*, number 1488 in Lecture Notes in Mathematics. Springer Verlag, 1991.

[Joh77] Peter Johnstone. *Topos Theory.* Number 10 in London Mathematical Society Monographs. Academic Press, 1977.

[Joh82] Peter Johnstone. *Stone Spaces.* Number 3 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1982.

[Joh84] Peter Johnstone. Open locales and exponentiation. *Contemporary Mathematics*, 30:84–116, 1984.

[JR84] William Julian and Fred Richman. A uniformly continuous function on [0,1] that is everywhere different from its infimum. *Pacific Journal of Mathematics*, 111:333–340, 1984.

[JT84] André Joyal and Myles Tierney. *An Extension of the Galois Theory of Grothendieck*, volume 309 of *Memoirs.* American Mathematical Society, 1984.

[Kau80] Edgar Kaucher. Interval analysis in the extended interval space IR. In Götz Alefeld and Rolf Grigorieff, editors, *Fundamentals of Numerical Computation*, volume 2 of *Computing. Supplementum.* Springer-Verlag, 1980.

[Kea96] Baker Kearfott. Interval computations: Introduction, uses and resources. *Euromath Bulletin*, 2(1):95–112, 1996.

[Kel86] Max Kelly. A survey of totality in ordinary and enriched categories. *Cahiers de Géometrie et Topologie Differentielle*, 27:109–132, 1986.

[Kle45] Stephen Kleene. On the interpretation of intuitionistic number theory. *Journal of Symbolic Logic*, 10:109–124, 1945.

[KNZ96] Vladik Kreinovich, Vyacheslav Nesterov, and Nina Zheludeva. Interval methods that are guaranteed to underestimate (and the resulting new justification of Kaucher arithmetic). *Reliable Computing*, 2(2):119–124, 1996.

[Koc06] Anders Kock. *Synthetic Differential Geometry.* Number 333 in London Mathematical Society Lecture Note Series. Cambridge University Press, second edition, 2006.

[Lak95] Anatoly Lakeyev. Linear algebraic equations in Kaucher arithmetic. In Vladik Kreinovich, editor, *Applications of Interval Computations (APIC'95)*, 1995. supplement to *Reliable Computing.*

[Lam07] Branimir Lambov. RealLib: An efficient implementation of exact real arithmetic. *Mathematical Structures in Computer Science*, 17(1):81–98, 2007. `www.brics.dk/~barnie/RealLib/`.

[Moo66] Ramon Moore. *Interval Analysis.* Prentice Hall, 1966.

[Mül01] Norbert Müller. The iRRAM: Exact arithmetic in C++. In Jens Blanck, Vasco Brattka, and Peter Hertling, editors, *Computability and Complexity in Analysis: 4th International Workshop, CCA 2000 Swansea, UK, September 17, 2000, Selected Papers*, number 2064 in Lecture Notes in Computer Science. Springer-Verlag, 2001.

[Pal05] Eric Palmgren. Continuity on the real line and in formal spaces. In Laura Crosilla and Peter Schuster, editors, *From Sets and Types to Topology and Analysis: Towards Practicable Foundations of Constructive Mathematics*, Oxford Logic Guides. Oxford University Press, 2005.

[Pea97] Giuseppe Peano. Studii di logica matematica. *Atti della Reale Accademia di Torino*, 32:565–583, 1897. Reprinted in Peano, *Opere Scelte*, Cremonese, 1953, vol. 2, pp. 201–217, and (in English) in Kennedy, Hubert, *Selected Works of Giuseppe Peano*, Toronto University Press, 1973, pp 190–205.

[Plo77] Gordon Plotkin. LCF considered as a programming language. *Theoretical Computer Science*, 5:223–255, 1977.

[Rob66] Abraham Robinson. *Non-standard Analysis*. North-Holland, 1966. Revised edition, 1996, published by Princeton University Press.

[Rog92] Hartley Rogers. *Theory of Recursive Functions and Effective Computability*. MIT Press, third edition, 1992.

[Ros86] Giuseppe Rosolini. *Continuity and Effectiveness in Topoi*. PhD thesis, University of Oxford, 1986.

[Ros01] Frank Rosemeier. A constructive approach to Conway's theory of games. In Peter Schuster, Ulrich Berger, and Horst Osswald, editors, *Reuniting the Antipodes: Constructive and Nonstandard Views of the Continuum*. Springer-Verlag, 2001.

[Sch03] Peter Schuster. Unique existence, approximate solutions, and countable choice. *Theoretical Computer Science*, 305:433–455, 2003.

[Sch05] Peter Schuster. What is continuity, constructively? *Journal of Universal Computer Science*, 11(12):2076–2085, 2005.

[Sco72a] Dana Scott. Continuous lattices. In Bill Lawvere, editor, *Toposes, Algebraic Geometry and Logic*, number 274 in Lecture Notes in Mathematics. Springer-Verlag, 1972.

[Sco72b] Dana Scott. Lattice theory, data types and semantics. In Randall Rustin, editor, *Formal Semantics of Programming Languages*. Prentice-Hall, 1972.

[Spi07] Bas Spitters. Located and overt sublocales. `arxiv.org/abs/math/0703561`, 2007.

[Ste85] David Stevenson. Binary floating-point arithmetic. *ANSI/IEEE Standard*, 754, 1985. Revised, 2008.

[Sto83] Otto Stolz. Zur Geometrie der Alten, insbesondere über ein Axiom des Archimedes. *Mathematische Annalen*, 22(4):504–519, 1883.

[Str80] Ross Street. Cosmoi of internal categories. *Transactions of the American Mathematical Society*, 258:271–318, 1980.

[SW78] Ross Street and Robert Walters. Yoneda structures on 2-categories. *Journal of Algebra*, 50:350–379, 1978.

[Tay91] Paul Taylor. The fixed point property in synthetic domain theory. In Gilles Kahn, editor, *Logic in Computer Science 6*. IEEE, 1991.

[Tay99] Paul Taylor. *Practical Foundations of Mathematics*. Number 59 in Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.

[Thi97] Hayo Thielecke. *Categorical Structure of Continuation Passing Style*. PhD thesis, University of Edinburgh, 1997. ECS-LFCS-97-376.

[Tho80] Walter Tholen. A note on total categories. *Bulletin of the Australian Mathematical Society*, 21:169–173, 1980.

[TvD88] Anne Sjerp Troelstra and Dirk van Dalen. *Constructivism in Mathematics, an Introduction*. Number 121 and 123 in Studies in Logic and the Foundations of Mathematics. North-Holland, 1988.

[Ver94] Japie Vermeulen. Proper maps of locales. *Journal of Pure and Applied Algebra*, 92:79–107, 1994.

[Waa05] Frank Waaldijk. On the foundations of constructive mathematics – especially in relation to the theory of continuous functions. *Foundations of Science*, 10(3):249–324, 2005.

[Wei00] Klaus Weihrauch. *Computable Analysis*. Springer-Verlag, Berlin, 2000.

[Woo82] Richard Wood. Some remarks on total categories. *Journal of Algebra*, 75:538–545, 1982.

The papers on abstract Stone duality may be obtained from

`www.Paul Taylor.EU/ASD`

[O] Paul Taylor, Foundations *for* Computable Topology. in Giovanni Sommaruga (ed.), *Foundational Theories of Mathematics*, Kluwer 2009.

[A] Paul Taylor, Sober spaces and continuations. *Theory and Applications of Categories*, 10(12):248–299, 2002.

[B] Paul Taylor, Subspaces in abstract Stone duality. *Theory and Applications of Categories*, 10(13):300–366, 2002.

[C] Paul Taylor, Geometric and higher order logic using abstract Stone duality. *Theory and Applications of Categories*, 7(15):284–338, 2000.

[D] Paul Taylor, Non-Artin gluing in recursion theory and lifting in abstract Stone duality. 2000.

[E] Paul Taylor, Inside every model of Abstract Stone Duality lies an Arithmetic Universe. *Electronic Notes in Theoretical Computer Science* **122** (2005) 247-296, Elsevier.

[F] Paul Taylor, Scott domains in abstract Stone duality. March 2002.

[G–] Paul Taylor, Local compactness and the Baire category theorem in abstract Stone duality. *Electronic Notes in Theoretical Computer Science* **69**, Elsevier, 2003.

[G] Paul Taylor, Computably based locally compact spaces. *Logical Methods in Computer Science*, **2** (2006) 1–70.

[H–] Paul Taylor, An elementary theory of the category of locally compact locales. APPSEM Workshop, Nottingham, March 2003.

[H] Paul Taylor, An elementary theory of various categories of spaces and locales. November 2004.

[J] Paul Taylor, A λ-calculus for real analysis. Computability and Complexity in Analysis, Kyoto, August 2005.

[K] Paul Taylor, Interval analysis without intervals. February 2006.

[L] Paul Taylor, Tychonov's theorem in abstract Stone duality. September 2004.

[M] Paul Taylor, Cartesian closed categories with subspaces. 2009.

This paper results from a collaboration that began with Paul Taylor's visit to Ljubljana in November 2004. A preliminary version of this work was presented at *Computability and Complexity in Analysis* in Kyoto on 28 August 2005, and we are grateful to Peter Hertling and the CCA programme committee for the indulgence of allowing us to occupy altogether 80 pages of their proceedings. We would also like to thank Vasco Brattka, Douglas Bridges, Thierry Coquand, Fer-Jan de Vries, Peter Johnstone, Vladek Kreinovich, Russell O'Connor, Andrea Schalk, Peter Schuster, Alex Simpson, Bas Spitters, Chris Stone, Maarten van Emden and Graham White for their helpful comments, and the anonymous referee for a most professional report.

CORRESPONDENCE (by email only, please):
Andrej Bauer, Department of Mathematics and Physics, University of Ljubljana.
Paul Taylor.
`Andrej.Bauer@andrej.com`   `pt08@PaulTaylor.EU`