Chapter 11 Explanation

Alexander Grabanski - ajg137@case.edu

July 6, 2017

1 Dedekind Reals

The reals, as constructed by two-sided Dedekind cuts in the HoTT book, are nothing new for constructive mathematics. Fundamentally, the construction of a (two-sided) Dedekind real relies on defining a *lower cut* and an *upper cut*, which are both disjoint inhabited subsets of the rationals such that they are also rounded (for every rational in each, there is always a rational closer to the edge of the cut, as are there rationals further from the edge) and located ($(q < r)$ implies that q is in the lower cut, or r is in the upper cut).

Porting this common definition over to HoTT works just fine if we interpret "subset" as "function targeting the space of mere propositions", but we can only hope to do something like that in the presence of impredicativity or LEM (which'd collapse all propositions to just **2**). Ignoring this impredicativity, the Dedekind reals have some nice properties – in particular, if we define a Cauchy approximation as a function of type $\mathbb{Q}_+ \to \mathbb{R}_D$ satisfying the predicate:

$$isCauchyApx(x) : \forall(\delta, \epsilon : \mathbb{Q}_+). \quad |x_\delta - x_\epsilon| < \delta + \epsilon$$

which is a sharpening of the classical concept of a Cauchy sequence which removes the inner "there exists a natural number such that..." to concern ourselves only with the closeness of approximations from the start. With this definition of a Cauchy approximation, $\mathbb{R}_D$ is Cauchy-complete, in that every such approximation has a limit in $\mathbb{R}_D$. It also is an Archimedian totally-ordered field, and every such field embeds in it, so what more could we want?

The big problem with Dedekind cuts, aside from impredicativity of some constructions, is that they are relatively opaque to constructive methods. Everywhere in their definition, we can see propositionally-truncated existential quantifiers and "or" clauses. Moreover, the description of the Dedekind reals does not naturally lend itself to a computational point-of-view, since deciding whether a real is in an certain side of the cut reduces to deciding whether a proposition is inhabited, which is exactly what LEM is often needed for. We need something better, hopefully leveraging unique construction methods available to HoTT.

2 Cauchy Reals

2.1 A Failed Attempt

There is one other commonly-used classical definition of the real numbers: equivalence classes of Cauchy sequences of rationals. Here, we could hope to provide a higher inductive type $Cauchy$ with a single point constructor:

$$lim : CauchyApx \to Cauchy$$

Where we define $CauchyApx$ by:

$$\sum_{x:\mathbb{Q}_+ \to \mathbb{Q}} isCauchyApx(x)$$

where $isCauchyApx$ is an analogous version of the predicate for "Cauchy Approximations" defined in the case of the Dedekind Reals, but restricted to sequences of rationals. Then, we would have the burden of defining a suitable notion of equality of such sequences, which we could achieve through:

$$eq_{Cauchy} : \prod_{x,y:CauchyApx} (\prod_{\epsilon:\mathbb{Q}_+} |pr_1(x)(\epsilon/2) - pr_1(y)(\epsilon/2)| < \epsilon) \to lim(x) =_{Cauchy} lim(y)$$

which aligns with the usual definitions, and we could construct basic operations (addition/multiplication) by lifting operations on the rationals pointwise to sequences. We could also easily construct notions of order, and prove that we have an ordered field.

But then, if we attempt to prove that this simple version of the Cauchy reals is Cauchy-complete, we run into a problem: Suppose that we have a Cauchy approximation $A : \mathbb{Q}_+ \to Cauchy$ of $Cauchy$s. Then, to extract a limit of A, we would need to construct a sequence of rationals converging to it. If we interpreted all elements of $Cauchy$ as raw $CauchyApx$es, doing this would be very straightforward – we could simply use $A(\epsilon/3)(\epsilon/3)$ as our sequence. However, the "elements" of $Cauchy$ are *equivalence classes* of Cauchy approximations, and so we would need to invoke the Axiom of Countable Choice to make an appropriate choice for every $\epsilon : \mathbb{Q}_+$, which is not provided in constructive HoTT.

While the above construction is somewhat simple, and mirrors classical developments, a construction of the reals which isn't Cauchy-Complete doesn't seem worthy of the name "Cauchy". In particular, many classic arguments of basic real analysis require the completeness of the reals to go through. Luckily, there is a better way to define a higher-inductive(-inductive) type representing real numbers which maintains the spirit of the above definition, but *is* Cauchy-complete. This improved definition is what the book calls $\mathbb{R}_C$.

2.2 Cauchy Reals in HoTT

In the failed attempt above, we ran into trouble because a Cauchy approximation of $Cauchy$s did not have a canonical representation using the only point constructor of

$Cauchy$. So, in the spirit of making a definition to make the theorems work, let's widen the point constructor to allow Cauchy sequences of Cauchy reals to define new Cauchy reals, like so:

$$lim : CauchyApx \to \mathbb{R}_C$$

where

$$CauchyApx \equiv \sum_{x:\mathbb{Q}_+ \to \mathbb{R}_C} isCauchyApx(x)$$

but this is a circular definition with no base case, and so we should throw in the rationals using a point constructor:

$$rat : \mathbb{Q} \to \mathbb{R}_C$$

These seem like a reasonable starting point to express the real numbers as freely-generated by rationals and the limit operation, or in other words, a Cauchy completion of the rationals. However, the usual definition $isCauchyApx$ cannot be defined in terms of these constructors, because in particular, we would need to express notions of subtraction, absolute value, and order to define "closeness", but doing so explicitly in terms of underlying sequences of rationals would run up against Countable Choice for the simple case of a Cauchy approximation of Cauchy approximations of rationals. It seems like we're stuck yet again.

The clever insight made by the authors of the HoTT book was that the notion of "closeness" may be defined *simultaneously* with $\mathbb{R}_C$, at the expense of replacing a higher inductive formulation with a so-called higher inductive-inductive formulation. In it, with $\mathbb{R}_C$, simultaneously and inductively define a relation $\sim : \mathbb{Q}_+ \times \mathbb{R}_C \times \mathbb{R}_C \to \mathcal{U}$, utilizing the syntactic sugar $x \sim_\epsilon y$ for $\sim (\epsilon, x, y)$. Intuitively, this relation should express $|x - y| < rat(\epsilon)$, which the book later proves as an equivalence in Theorem 11.3.44. For now, it's our job to express how the different pairings of constructors of $\mathbb{R}_C$ interact with $\sim_\epsilon$ to give the relation meaning.

Starting with the easy case of two rationals, we already have notions of subtraction and order, and so we can provide a constructor:

$$\prod_{q,r:\mathbb{Q}} \prod_{\epsilon:\mathbb{Q}_+} (|q - r| < \epsilon) \to rat(q) \sim_\epsilon rat(r)$$

The more complicated cases involve pairings of limits with rationals and other limits. Suppose that we have $a \equiv rat(q)$ and $b \equiv lim(y)$. To ease reasoning about what this case should look like, consider the symbol L as the limit of y (though we haven't defined that notion yet!). Intuitively, for any $\delta : \mathbb{Q}_+$, $y_\delta \sim_\delta L$, and so we may consider y_δ as lying in the open interval $(L - \delta, L + \delta)$. We want to determine if $(q - \epsilon, q + \epsilon)$ contains L based purely on y_δ. For one worst-case, $q + \epsilon$ could be L or greater, in which case we could only guarantee that y_δ would be less than $q + \epsilon$ if $rat(q) \sim_{\epsilon-\delta} y_\delta$. The other, symmetric worst case yields the same condition, and so with this motivation, we provide a constructor of type:

$$\prod_{q:\mathbb{Q}} \prod_{y:CauchyApx} \prod_{\epsilon,\delta:\mathbb{Q}_+} (rat(q) \sim_{\epsilon-\delta} y_\delta) \to (rat(q) \sim_\epsilon lim(y))$$

We also define the similar, symmetric case for where the left-hand side of the relation is a limit, and the right-hand side is a rational.

The case of two limits is analogous to the situation above, and the intuitive argument essentially carries through if we replace q with an approximation to a real within some error bounded by $\eta : \mathbb{Q}_+$. This yields a final constructor for $\sim_\epsilon$, through:

$$\prod_{x,y:CauchyApx} \prod_{\epsilon,\delta,\eta:\mathbb{Q}_+} (x_\eta \sim_{\epsilon-\delta-\eta} y_\delta) \to (lim(x) \sim_\epsilon lim(y))$$

Now that "closeness" is on a firm foundation with $\sim_\epsilon$, the definition of $isCauchyApx$ falls out naturally:

$$isCauchyApx(x) :\equiv \prod_{\epsilon,\delta:\mathbb{Q}_+} x_\delta \sim_{\delta+\epsilon} x_\epsilon$$

Now, for the higher constructors, first demand that every $x \sim_\epsilon y$ is a proposition by providing an equality constructor for every two elements. Then, all that remains is to re-introduce the "equivalence class" part of the usual definition of Cauchy reals by providing a higher constructor in $\mathbb{R}_C$ to say that reals that are ϵ-close for every ϵ are equal:

$$eq_{\mathbb{R}_C} : \prod_{u,v:\mathbb{R}_C} \forall(\epsilon : \mathbb{Q}_+).u \sim_\epsilon v \to u =_{\mathbb{R}_C} v$$

The book gives a compact listing of these constructors on p383. Some of them use $\forall$ in place of $\prod$, but the two coincide in usage since $\sim$ is a mere relation.

2.3 Induction Principles

Of course, we can now easily supply a limiting value for any Cauchy approximation by rationals, or any Cauchy approximation by reals that we've already defined, but datatypes would be boring if the only thing we could do is to construct elements of them. If we're ever going to hope to be able to prove "the reals form a complete archimedean ordered field", we'll need to define how some basic operations act on the constructors of the inductively-defined pair $(\mathbb{R}_C, \sim)$. So we'll need an induction principle on the pair, which means that we'll first need to figure out what type families work for the inductive-inductive structure.

$\mathbb{R}_C$ is just a type, so elements of the type are all we can hope to predicate a type family on. Let

$$A : \mathbb{R}_C \to \mathcal{U}$$

Then, on the other hand, $\sim$ is best viewed as an inductively-defined *family* of types. To define the output of a type family predicated on it, we will need all of the information that went into construction of a particular (type-valued) instance of $\sim$ (for $x \sim_\epsilon y$, x, y, and ϵ), but we'll also need both the particular element of $x \sim_\epsilon y$ *and* the particular choices of elements of $A(x)$ and $A(y)$. The last two pieces of data come from the fact that the fibration induced by A is one part of the larger "inductive-inductive fibration" we're defining. As a thought experiment, imagine if the type family over relations $\sim$ had type:

$$B : \prod_{x,y:\mathbb{R}_C} \prod_{\epsilon:\mathbb{Q}_+} (x \sim_\epsilon y) \to \mathcal{U}$$

Then, the fibration we could expect to get from this would take quadruples $(x, y, \epsilon, p, B(x, y, \epsilon, p))$ to (x, y, ϵ, p). The path-lifting property of fibrations would lift paths in those quadruples to paths in the total space, but there's something just not quite right about this – when we lifted paths in $\mathbb{R}_C$ before as part of this "inductive-inductive fibration", we got paths in the space $\sum_{x:\mathbb{R}_C} A(x)$. We should be consistent about how paths in certain types behave under the lift defined by the fibration as a whole. As a result, we should change the quadruple the tuple $(x, y, u : A(x), v : A(y), \epsilon, p, B(x, y, \epsilon, p))$, and still project to (x, y, ϵ, p). Then path-lifting may be straightforwardly imposed on the extra parameters (in the positions taken by u, v) through the lifting induced by A.

So, in analogy with how (actual!) fibrations of type $\sum_{x:A} P(x) \to A$ induce "fibrations" (in the sense of HoTT) typed $\prod_{x:A} P(x)$, the correct type for B is then:

$$B : \prod_{x,y:\mathbb{R}_C} A(x) \to A(y) \to \prod_{\epsilon:\mathbb{Q}_+} (x \sim_\epsilon y) \to \mathcal{U}$$

The book shortens this B to $(x, u) \frown_\epsilon (y, v)$, which omits the dependence on $(x \sim_\epsilon y)$ (which is a mere proposition anyway – one just needs to exist anywhere we have such an object), and then further shortens this to $u \frown_\epsilon v$ for when (x, y) are clear based on the context. Due to propositional truncation of $\sim$, all of the $\frown$-derived types will wind up inheriting the truncation anyway, so we are justified in writing it infix and thinking of it as a mere relation.

The induction principle on $(\mathbb{R}_C, \sim)$ is complicated due to the somewhat lengthy list of constructors. It appears on p384 in full, but there are some important things to be emphasized about it. The book defines a *dependent Cauchy approximation over x* as an element of

$$dependCauchy(x) :\equiv \sum_{a:\prod_{\epsilon:\mathbb{Q}_+} A(x_\epsilon)} \forall(\delta, \epsilon : \mathbb{Q}_+)(x_\delta, a_\delta) \frown_{\delta+\epsilon} (x_\epsilon, a_\epsilon)$$

As seems to be common in the chapter, the book often refers to elements of this type by its first projection, taking the purely-propositional part to be separate and given. Then, the $\mathbb{R}_C$ part of the induction principle simply demands that every rational-derived real gives some element of A by itself (called f_q for $q : \mathbb{Q}$), and any $(x, a : dependCauchy(x))$ defines something in $A(lim(x))$ for the action on the real given by $lim(x)$ (written $f_{x,a}$). The $\sim$ part of the induction principle may be derived from the constructors by replacing all instances of $\sim$ with $\frown$, all appearances of rational numbers with (q, f_q) pairs, all Cauchy approximations with (Cauchy approximation, dependent Cauchy approximation) pairs, and all references to $lim(x)$ with $(lim(x), f_{x,a})$ for a the dependent Cauchy approximation over x.

Also crucially, the $eq_{\mathbb{R}_C}$ higher constructor of $\mathbb{R}_C$ needs to be respected.under transport in A. Showing this is usually a pain.

Once the behavior of a desired function has been defined on rationals and limits, and the relation $\frown$ has been shown to be compatible with what we expect from $\sim$, the induction principle yields two functions:

$$f : \prod_{x:\mathbb{R}_C} A(x)$$

$$g : \prod_{x,y:\mathbb{R}_C} \prod_{\epsilon:\mathbb{Q}_+} \prod_{\zeta:x\sim_\epsilon y} B(x, y, f(x), f(y), \epsilon, \zeta)$$

It's clear that the type of g is quite long, but it's necessary to derive Cauchy approximations over some $x : CauchyApx$ consistent with the induction. (Where the meaning of "consistent with the induction" is specified entirely by the choices made during the induction on $\frown$) All we need to know is that the function f satisfies:

$$f(rat(q)) :\equiv f_q$$

and

$$f(lim(x)) :\equiv f_{x,a}$$

for a the dependent Cauchy approximation over x consistent with the induction.

The length of this induction principle makes it frankly awful to work with. Luckily, we don't always need its full power – specializing B to always return $\mathbf{1}$ yields a principle called $\mathbb{R}_C$-induction, which only demands that we specify the behavior of f_x and $f_{x,a}$ and prove that $eq_{\mathbb{R}_C}$ preserves it under transport in A, where now we may simply take $a : \prod_{\epsilon:\mathbb{Q}_+}$ since the second projection in $dependCauchy(x)$ is now $*$. The g we obtained previously is now entirely trivial.

From $\mathbb{R}_C$-induction, it follows immediately (Lemma 11.3.8) that $x \sim_\epsilon x$ for all x and ϵ, since all rationals are obviously ϵ-close to themselves for any ϵ, and limits of Cauchy approximation are ϵ-close to themselves for any ϵ since $x_{\epsilon/3} \sim_{\epsilon-(\epsilon/3)-(\epsilon/3)} x_{\epsilon/3}$ follows from the inductive hypothesis on terms x_δ of Cauchy approximations, which directly implies $lim(x) \sim_\epsilon lim(x)$ by the limit-limit constructor of $\sim$. Then, this implies (through $eq_{\mathbb{R}_C}$) that $\mathbb{R}_C$ is a set, since we have a mere relation implying equality as in the Lemma in Chapter 7 which lead to the proof of Hedberg's Theorem.

Through $\mathbb{R}_C$-induction, the book also proves that the Cauchy reals are merely limit points of $\mathbb{R}_C$ (Lemma 11.3.10) and that any function on Cauchy approximations which preserves limits defines a function on Cauchy reals. The book also (very briefly) considers $\sim$-induction by taking $A \equiv \mathbf{1}$, but only to prove that $\sim$ is symmetric, which we could have done away with by demanding it from the outset in the definition of our higher inductive-inductive type.

2.4 Getting Somewhere: Recursion Principles

The major barrier that the book identifies to using the full induction principle is in the requirements imposed by the higher constructor $eq_{\mathbb{R}_C}$ w.r.t. transport, and so the book just throws away the dependent information in A and B create a recursion principle for $(\mathbb{R}_C, \sim)$. The complicated condition on $eq_{\mathbb{R}_C}$ turns into the simpler demand that the relations $\frown$

are separated, meaning that $\frown_\epsilon$ for all ϵ is enough to prove an equality. This recursion principle is immediately employed to prove the most important lemma for establishing the basic algebraic operations on $\mathbb{R}_C$, and is listed on p388.

With F taken to be either $\mathbb{Q}$ or $\mathbb{R}_C$, a function $f : F \to \mathbb{R}_C$ is called *Lipschitz* if there is a *Lipschitz constant* $L : \mathbb{Q}_+$ such that $\prod_{\epsilon:\mathbb{Q}_+} \prod_{u,v:F}(u \sim_\epsilon v) \to f(u) \sim_{L*\epsilon} f(v)$ for $(u \sim_\epsilon v)$ replaced with "$|u - v| < \epsilon$" in the case of $F \equiv \mathbb{Q}$. Intuitively, this says that everywhere, the distance between two points may only increase up to a constant magnification factor when passed through f. This intuition is why the case where $L < 1$ leads to f being called a "contraction map".

Using a straighforward recursion, Lipschitz functions on $\mathbb{Q}$ lift to Lipschitz functions on $\mathbb{R}_C$ which agree on the rationals. The only trick is to define an appropriate incidence relation $\frown$, keeping in mind that this relation is defined on the *codomain* of the function we want to extract from the recursion. Using $u \frown_\epsilon v :\equiv u \sim_{L*\epsilon} v$ makes the most sense, since our goal is to show that the resulting function we are defining is Lipschitz.

The proof is straightforward once we define our target function $\bar{f}$ to agree on rationals with the original, and define its action on limits by taking

$$\bar{f}(lim(x)) :\equiv lim(\lambda\epsilon.\bar{f}(x_{\epsilon/L})$$

which works as a Cauchy approximation since we may suppose that we have a dependent Cauchy approximation over such an x as the IH, which means that we have

$$\bar{f}(x_\delta) \sim_{L*\delta+L*\epsilon} \bar{f}(x_\epsilon)$$

for any δ and ϵ, and so dividing through dividing all rational indices through by L yields an element of $isCauchyApx(\bar{f}(lim(x)))$.

On the $\frown$ relation, in every constructor case involving limits, the dependent Cauchy approximations in the IHs always give enough information to just pass them back to the same constructor to obtain a version multiplied by L. The $\frown$ relation also directly inherits separatedness from $\sim$.

Using this lemma, we can lift unary negation from the rationals to the reals (Lipschitz constant 1) [and more generally, multiplication by any rational], but we can also lift *squaring*, which is for the construction of multiplication, and we also can lift multiplicative inverses. The trick to the latter two is to break up parts of the domain into nested closed intervals and then prove that the functions are Lipschitz on those intervals (deriving Lipschitz constants on these is an exercise in basic algebra). Then, using the fact that $\mathbb{R}_C$ is a set, we can use the nice properties of the category of sets to construct the whole function from a direct limit of the functions on the closed intervals. Of course, this means we need a notion of "closed intervals" first.

As an intermediate step, the book proves a nice lemma with an awful proof characterizing a relation which turns out to be equivalent to $x \sim_\epsilon y$ by mere existence of the other variable(s) in the constructor(s) (if any) satisfying the properties of their respective domains. It also proves (together with the equivalence proof) the triangle inequality:

$$u \sim_\epsilon v \to v \sim_\delta w \to u \sim_{\epsilon+\delta} w$$

that ϵ-close values are also δ-close for when δ is bigger (monotonicity), and that the bound defined by $\sim_\epsilon$ is an open set – we can always merely find a rational δ smaller than ϵ such that two values are δ-close. The proof is done by a double-recursion, which leads to an explosion in the number of proof-cases. Nevertheless, it is provable with care.

For binary operations on the reals, Lemma 11.3.40 leverages the lifting of Lipschitz functions to prove that maps $f : \mathbb{Q} \times \mathbb{Q} \to \mathbb{Q}$ which are contraction maps in each argument (with the other fixed) lift to the reals, which immediately implies that $-$, $+$, max, and min are all definable. These, in turn, imply that we may construct the non-strict ordering on the reals (which means we may define a "closed interval") and multiplication by $a * b :\equiv ((a + b)^2 - a^2 - b^2)/2$.

The proof of that lemma, once again, falls out (for the most part) of an appropriate definition of A, taken to be the set of all real-valued contraction maps, and an appropriate definition of $\frown_\epsilon$, taken to be the relation "values of functions in A for all arguments are within ϵ. For rationals, the definition of $\bar{f}$ is taken to be the standard lifting of a Lipschitz function, and for reals, the limit-swizzling trick as in the proof for Lipschitz-lifting is employed again. Preservation of the incidence relations is just as easy to prove as it was for the Lipschitz-lifting lemma.

After that, a definition for strict ordering $<$ is given by demanding

$$(u < v) :\equiv \exists q, r : \mathbb{Q} \quad (u \leq q) \wedge (q < r) \wedge (v \geq r)$$

which immediately yields the Archimedian principle through merely averaging q and r (taken here to mean "mere existence of a rational strictly between any two reals").

Then, a few straightforward lemmas on orderings lead to the proof that $\sim_\epsilon$ is exactly what we wanted it to be –

$$x \sim_\epsilon y \simeq (|x - y| < \epsilon)$$

This completes the description of the basic structure of the Cauchy reals, since we can show that they are a field by simply showing that our lifted operations inherit their properties from the rationals, they are Cauchy-complete by construction, and the Archimedian principle follows directly from our definition of the strict ordering.

2.5 Under Classical Assumptions...

Since we have the expected structure on Cauchy reals, they embed into the Dedekind reals. But we get a stronger statement under LEM or ACC – the two notions coincide! Both of these fundamentally come down to a method for extracting a Cauchy approximation from the specification of a Dedekind real, and showing that the real is merely the limit of the approximation. The case of LEM is straightforward – starting with the inhabitation condition for a Dedekind cut, we merely have something in the lower cut, and in the upper cut. From there, a computer scientist's bread-and-butter, the binary search, will yield a procedure for obtaining an arbitrarily-close approximation to the real number. The book defines a condition which works for countable choice as well as LEM, relying on locatedness of Dedekind cuts to derive a sequence instead.

As a result, we can trust the construction of the Cauchy reals, knowing that they reduce to Dedekind reals under classical assumptions.