Algebraic Real Analysis

PETER FREYD
pjf@upenn.edu

f

Special thanks to Mike Barr and Don von Osdol for editorial assistance and to the Executive Director of the FMS Foundation for making it all possible.

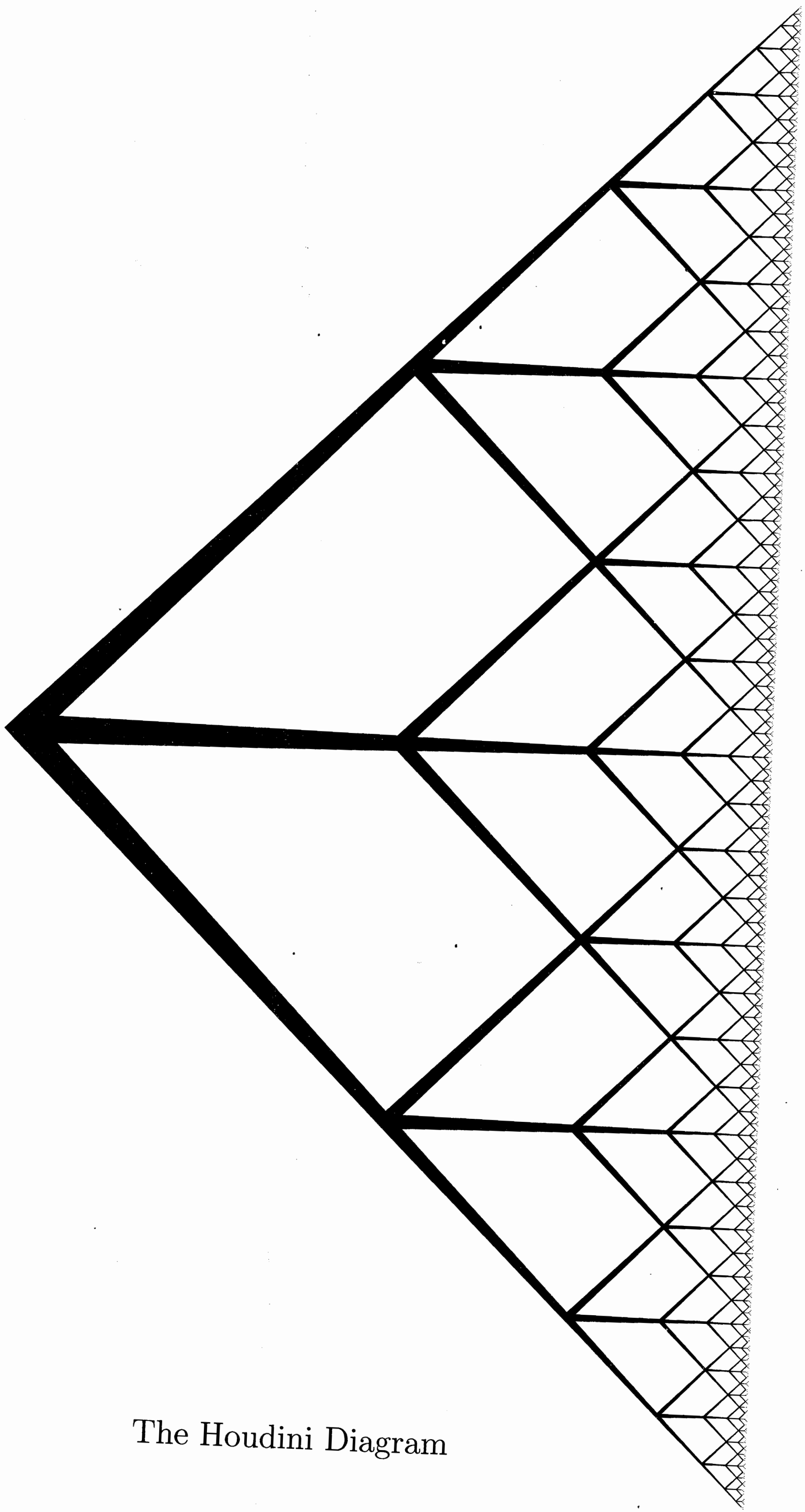
An earlier draft appeared in TAC (2008-07-02). This is the current revision. It is a reconstruction (with Niuniu Zhang's expert aid) of a printed document produced by a now-destroyed latex program. It is not yet the final version.

The earlier version can be found at http://www.tac.mta.ca/tac/ or by anonymous ftp from ftp://ftp.tac.mta.ca/pub/tac/html/volumes/20/10/20-10.{dvi,fps,pdf}. The current revision is available at

https://drive.google.com/drive/folders/1rPnHnUrm1yfon0MOMFDA2uOiT8yxmwjE?usp=sharing

Look for "analysis."

At the same address is a short paper intended as an introduction to this one. Look for "e-pi".



The Houdini Diagram

0. Introduction

The title is wishful thinking; there ought to be a subject that deserves the name "algebraic real analysis."

Herein is a possible beginning.

For reasons that can easily be considered abstruse we were led to the belief that the closed interval—not the entire real line—is the basic structure of interest. Before describing those abstruse reasons, a theorem:

Let G be a compact group and $\mathtt{I}$ the closed interval. (We will not say which closed interval; to do so would define it as a part of the reals, belying the view of the closed interval as the fundamental structure.) Let $\boldsymbol{\mathcal{C}(G)}$ be the set of continuous maps from G to $\mathtt{I}$. We wish to view this as an algebraic structure, where the word "algebra" is in the very general sense, something described by operations and equations. In the case at hand, the only operators that will be considered right now are the constants, "top" and "bottom," denoted $\top$ and $\bot$, and the binary operation of "midpointing," denoted $x|y$. (There are axioms that will define the notion of "closed midpoint algebra" but since the theorem is about specific examples they're not now needed.) $\mathcal{C}(G)$ inherits this algebraic structure in the usual way ($f|g$, for example, is the map that sends $\sigma \in G$ to $(f\sigma)\big|(g\sigma) \in \mathtt{I}$). We use the group structure on G to define an action of G on $\mathcal{C}(G)$, thus obtaining a representation of G on the group of automorphisms of the closed midpoint algebra. (Fortunately no knowledge of the axioms is necessary for the definition of automorphism—or even homomorphism.) Let $(\mathcal{C}(G), \mathtt{I})$ be the set of closed-midpoint–algebra homomorphisms from $\mathcal{C}(G)$ to $\mathtt{I}$. Again we obtain an action of G.

0.1 THEOREM: *There is a unique G–fixed-point in* $(\mathcal{C}(G), \mathtt{I})$

There is an equivalent way of stating this:

0.2 THEOREM: *There is a unique G-invariant homomorphism from the closed midpoint algebra $\mathcal{C}(G)$ to $\mathtt{I}$.*

This theorem is mostly von Neumann's: the unique G-invariant homomorphism is integration, that is, it is the map that sends $f : G \to \mathtt{I}$ to $\int f \, d\sigma$. But it is not entirely von Neumann's: we have just characterized integration on compact groups without a single mention of semiquations or limits (see Section 40, p91–94, for proofs). The only non-algebraic notion that appeared was at the very beginning in the definition of $\mathcal{C}(G)$ as the set of *continuous* maps (in Section 11, p30–31, we will obtain a totally algebraic definition).

The fact that we are stating this theorem for $\mathtt{I}$ and not the reals, $\mathbb{R}$, is critical. Consider the special case when G is the one-element group; the theorem says that the identity map on $\mathtt{I}$ is the only midpoint-preserving endomorphism that fixes $\top$ and $\bot$ (we said that the theorem is mostly von Neumann's; this part is not, see Section 4, p14–17). It actually suffices to assume that the endomorphism fixes any two points (Lemma 13.1, p32) but with the axiom of choice and a standard rational Hamel-basis argument we can find $2^{2^{\aleph_0}}$ counterexamples for this assertion if $\mathtt{I}$ is replaced with $\mathbb{R}$ (even when the number of designated fixed-points is not just two but countably infinite).

We do not need a group structure or even von Neumann to make the point. Consider this remarkably simple characterization of definite integration of continuous maps from interval to interval.

$$\fint \top \, \mathrm{d}x = \top \qquad\qquad \fint \bot \, \mathrm{d}x = \bot$$

$$\fint f(x) \mid g(x) \, \mathrm{d}x \; = \; \fint f(x) \, \mathrm{d}x \;\mid\; \fint g(x) \, \mathrm{d}x$$

$$\fint f(x) \, \mathrm{d}x \; = \; \fint f(\bot|x) \, \mathrm{d}x \;\mid\; \fint f(\top|x) \, \mathrm{d}x \qquad [1]$$

No semiquations. No limits. The first three equations say just that integration is a homomorphism of closed midpoint algebras. The fourth equation says that the mean value of a function on $\mathtt{I}$ is the midpoint of the two mean-values of the function on the two halves of $\mathtt{I}$.

The fourth equation—as any numerical or theoretical computer scientist will tell you—is a "fixed-point characterization." When Church proved that his and Gödel's notion of computability were coextensive he used the fact that all computation can be reduced to finding fixed-points. (The word "point" here is traditional but misleading. The fixed-point under consideration here is, as it was for Church, an operator on functions—rather far removed from the public notion of point.)

If we seek a fixed-point of an operator the first thing to try, of course, is to iterate the operator on some starting point and to hope that the iterations converge. So let

$$\fint_0 f(x) \, \mathrm{d}x$$

denote an "initial approximation" operator, to wit, an arbitrary operator from $\mathcal{C}(\mathtt{I})$ to $\mathtt{I}$ that satisfies the first three equations. Define a sequence of operators, iteratively, as follows:

$$\fint_{n+1} f(x) \, \mathrm{d}x \; = \; \fint_n f(\bot|x) \, \mathrm{d}x \;\Big|\; \fint_n f(\top|x) \, \mathrm{d}x$$

where each new operator is to be considered an improvement of the previous. (One should verify that we automatically maintain the first three equations in each iteration.) Thus the phrase "fixed-point" here turns out to mean an operator so good that it can not be improved. (What is being asserted is that there is a unique such operator.) Wonderfully enough: *no matter what closed midpoint homomorphism is chosen as the initial approximation, the values of these operators are guaranteed to converge.*

If we take the initial approximation to be evaluation on $\bot$, that is, if we take

$$\fint_0 f(x) \, \mathrm{d}x \; = \; f(\bot)$$

then what we are saying turns out to be only that "left Riemann sums" work for integration. If we use $f(\top)$ for the initial approximation we obtain "right Riemann sums." For the "trapezoid rule" use the midpoint of these two initial operators, $f(\bot)|f(\top)$. For "Simpson's rule" use $\frac{1}{6}f(\bot) + \frac{2}{3}f(\bot|\top) + \frac{1}{6}f(\top)$. [2]

[1] There was an appendix for Latex macros, but the powers that be deemed such to be beneath the dignity of this journal. It appears here as Section 33 (p77–80).

[2] If this paper's title is to be taken seriously we will be obliged to give an algebraic description of the limits used in the previous paragraph. Here's one way: let $\mathtt{I}^{\mathbb{N}} = \prod_{\mathbb{N}} \mathtt{I}$ denote the closed midpoint algebra of all sequences in $\mathtt{I}$. The first step is to identify sequences that agree almost everywhere to obtain the quotient algebra $\mathtt{I}^{\mathbb{N}} \to A$. The latter will be shown—for entirely algebraic reasons—to have a closed midpoint homomorphism to $\mathtt{I}$ (Theorem 10.4, p28) and we could use any such homomorphism to define the sequential limits appearing in the previous paragraph. There is, of course, an obvious

1. *Diversion:* The Proximate Origins, or: Coalgebraic Real Analysis

The point of departure for this approach to analysis is the use of the closed interval as the fundamental structure; the reals are constructed therefrom. A pause to describe how I was prompted to explore such a view.

The community of theoretical computer scientists (in the European sense of the phrase) had found something called "initial-algebra" definitions of data types to be of great use. Such definitions typically tell one how inductive programs—and then recursive programs—are to be defined and executed.[3] It then became apparent that some types required a dual approach: something called "final-coalgebra definitions."[4] Such can tell one how "co-inductive" and "co-recursive" programs are to be defined and executed. (One must really resist here the temptation to say "co-defined" and "co-executed.")

Thus began a search for a final-coalgebra definition of that ancient data type, the reals. There is, actually, always a trivial answer to such a question: every object is automatically the final coalgebra of the functor constantly equal to that object. What was being sought was not just a functor with a final coalgebra isomorphic to the object in question but a functor that supplies its final coalgebra with the structure of interest. In 1999 an answer was found not for the reals but for the closed interval.[5] (To this date, no one has found a functor whose final coalgebra is *usefully* the reals.)

Consider, then, the category whose objects are sets with two distinguished points, denoted $\top$ and $\bot$ and whose maps are the functions that preserve $\top$ and $\bot$. Given a pair of objects, X and Y, we define their **ordered wedge**, denoted $\boldsymbol{X \vee Y}$, to be the result of identifying the top of X with the bottom of Y.[6] This construction can clearly be extended to the maps to obtain the "ordered-wedge functor." The closed interval can be defined as the final coalgebra of the functor that sends X to $X \vee X$. Let me explain.

First (borrowing from the topologists' construction of the ordinary wedge), $X \vee Y$ is taken as the subset of pairs, $\langle x, y\rangle$, in the product $X \times Y$ that satisfy the disjunctive condition: $x = \top$ *or* $y = \bot$.[7] A map, then, from X to $X \vee X$ may be construed as a pair of self-maps, whose values are denoted $\overset{\wedge}{\boldsymbol{x}}$ and $\overset{\vee}{\boldsymbol{x}}$, such that for all x either $\overset{\vee}{x} = \top$ or $\hat{x} = \bot$. The final coalgebra we seek is the terminal object in the category whose objects are these structures.

objection: we would be assigning limits to all sequences not just convergent ones; worse, the homomorphism would be not at all unique. Remarkably enough we can turn this inside out: an element in $\mathbf{I}^{\mathbb{N}}$ is convergent iff it is in the joint equalizer of all homomorphisms of the form $\mathbf{I}^{\mathbb{N}} \to A \to \mathbf{I}$. Put another way, $\lim a_n = b$ iff $h(\{a_n\}) = b$ whenever $h : \mathbf{I}^{\mathbb{N}} \to \mathbf{I}$ is a closed midpoint homomorphism that respects almost-everywhere equivalence. (See Section 11, p30–31, for an approach to this definition of derivatives that does not require the axiom of choice). This approach can be easily modified to supply limits of I-valued functions at points in arbitrary topological spaces. More interesting: it can be used to define derivatives. Let F be the set of all functions f from the standard interval $[-1,+1]$ to itself such that $|fx| \le |x|$. We will regard F as a closed midpoint algebra where the identity function is taken as $\top$ and the negation map as $\bot$. Now identify functions that name the same germ at 0 (that is, that agree on some neighborhood of 0) to obtain a quotient algebra $F \to A$. The joint equalizer of all homomorphisms of the form $F \to A \to [-1,+1]$ is precisely the set of functions differentiable at 0; the common values delivered by all such homomorphisms are the derivatives of those functions. That is, $f'(0) = b$ iff $H(f) = b$ whenever $H : F \to [-1,+1]$ is a closed midpoint homomorphism that depends only on the germs at 0 of its arguments. (Again, see Section 11, p30–31.)

[3] Let me brag here: I won a prize for one of a series of papers on this subject: Recursive types reduced to inductive types 5th Annual IEEE Symposium on Logic in Computer Science [LICS] (Philadelphia, PA. 1990), p408–507 *IEEE Comput. Soc. Press, Los Alamitos, CA*, 1990. The series culminated with Remarks on algebraically compact categories, Applications of categories in computer science (Durham, 1991), p95–106. *London Math. Soc. Lecture Note Ser.*, **177**, *Cambridge Univ. Press, Cambridge.* 1992.

[4] We don't actually need the general definitions, but for the record let T be an arbitrary endofunctor. A T-coalgebra is just a map $X \to TX$. A map between T-coalgebras is illustrated by the commutative diagram on the right. If the resulting category of T-coalgebras has a final object, it's called a final T-coalgebra.

$$\begin{array}{ccc} X & \to & TX \\ f\downarrow & & \downarrow Tf \\ Y & \to & TY \end{array}$$

[5] First announced in a note I posted on 22 December, 1999, `http://www.mta.ca/~cat-dist/1999/99-12`.

[6] The word "wedge" and its notation are borrowed from algebraic topology where $X \vee Y$ is the result of joining the (single) base-point of X to that of Y.

[7] Yes, $X \vee Y$ it a pushout: given $f : X \to Z$ and $g : Y \to Z$ such that $f\top = g\bot$ define $X \vee Y \to Z$ by sending $\langle x, y\rangle$ to fx `if` $y = \bot$ `else` gy.

To be formal, begin with the category whose objects are quintuples $\langle X, \bot, \top, {}^\wedge, {}^\vee \rangle$ where $\bot, \top \in X$, and ${}^\wedge, {}^\vee$ signify self-maps on X. The maps of the category are the functions that preserve the two constants and the two self-maps. Then cut down to the full subcategory of objects that satisfy the conditions:

$$\overset{\wedge}{\top} = \top = \overset{\vee}{\top}$$

$$\overset{\wedge}{\bot} = \bot = \overset{\vee}{\bot}$$

$$\forall_x \left[\hat{x} = \bot \quad \text{or} \quad \overset{\vee}{x} = \top \right]$$

$$\bot \neq \top \quad {}^{[8]}$$

We will call such a structure an **interval coalgebra**.[9]

I said that we will eventually construct the reals from $\mathbf{I}$. But if one already has the reals then one may chose $\bot < \top$ and define a coalgebra structure on $[\bot, \top]$ as

$$\overset{\vee}{x} = \min(2x - \bot, \top)$$

$$\hat{x} = \max(2x - \top, \bot)$$

Note that each of the two self-maps evenly expands a half interval to fill the entire interval—one the bottom half the other the top half. We will call them **zoom operators**. (By convention we will not say "zoom in" or "zoom out." All zooming herein is expansive, not contractive.)

The general definition of "final coalgebra" reduces—in this case—to the characterization of such a closed interval, $\mathbf{I}$, as the terminal object in this category.[10]

The general notion of "co-induction" reduces—in this case—to the fact that given any quintuple $\langle X, \bot, \top, {}^\wedge, {}^\vee \rangle$ satisfying the above-displayed conditions there is a unique $X \xrightarrow{f} \mathbf{I}$ such that $f(\bot) = \bot$, $f(\top) = \top$, $f(\overset{\vee}{x}) = (fx)^\vee$ and $f(\hat{x}) = (fx)^\wedge$. If $\mathbf{I}$ is taken as the unit interval, that is, if $\bot$ is taken as 0 and $\top$ as 1, then in the classical setting (and if one pays no attention to computational feasibility) a quick and dirty construction of f is to define the binary expansion of $f(x) \in \mathbf{I}$ by iterating (forever) the following procedure:

```
If x̌ = ⊤ then emit "1" and replace x with x̂
          else emit "0" and replace x with x̌
```

[8] I did not say "with a pair of distinguished points" above. What I said was "with *two* distinguished points." (Yes, I'm one of those who object to the phrase "these two things are the same.")

[9] The modal operations $\Diamond$ for *possibly* and $\Box$ for *necessarily* have received many formalizations but it is safe to say that no one allows simultaneously both $\Diamond\Phi \neq \top$ and $\Box\Phi \neq \bot$: "less than completely possible but somewhat necessary." (The coalgebra condition can be viewed here as a much weakened excluded middle: when the pair of unary operations are trivialized—that is, each taken to be the identity operation—then $\Diamond\Phi = \top$ *or* $\Box\Phi = \bot$ becomes just standard excluded middle.)

If we assume, for the moment, that $\top$ and $\bot$ are fixed points for $\Diamond$ and $\Box$ then we have an example of an interval coalgebra where $\Box\Phi = \overset{\wedge}{\Phi}$ and $\Diamond\Phi = \overset{\vee}{\Phi}$. The finality of $\mathbf{I}$ yields what may be considered truth values for sentences (*e.g.,* the truth value of $\bot|\top$ translates to "entirely possible but totally unnecessary" and a truth value greater than $\top|(\top|\bot)$ means "necessarily entirely possible").

The fixed-point conditions are not, in fact, appropriate—true does not imply necessarily true nor does possibly false imply false—but, fortunately, they're not needed: an easy corollary of the finality of $\mathbf{I}$ says that it suffices to assume the disjointness of the orbits of $\top$ and $\bot$ under the action of the two operators. If we work in a context in which the modal operations are monotonic (that is, when Φ implies Ψ it is the case that $\Box\Phi$ implies $\Box\Psi$ and $\Diamond\Phi$ implies $\Diamond\Psi$) it suffices to assume that $\Box\Phi$ implies Φ, that Φ implies $\Diamond\Phi$ and that $\Box^n\top$ never implies $\Diamond^n\bot$. If this last condition has never previously been stated it's only because no one ever thought of needing it.

The same treatment of modal operators holds when $\Diamond$ is interpreted as *tenable/conceivable/allowed/foreseeable* and $\Box$ as *certain/known/required/expected.* (Note that $\Diamond$ and $\Box$ are "De Morgan duals," that is, $\Box\Phi = \neg\Diamond(\neg\Phi)$ and $\Diamond\Phi = \neg\Box(\neg\Phi)$.) This topic will be much better discussed in the intuitionistic foundations considered in Section 29 (see [128], p64).

[10] If the case with $\bot = \top$ were allowed then the terminal object would be just the one-point set. (In some sense, then, the separation of $\top$ and $\bot$ requires no less than an entire continuum.)

Note that the symmetry on I is forced by its finality: if $\top$ and $\bot$ are interchanged and if $\top$- and $\bot$-zooming are interchanged the definition of interval coalgebra is maintained, hence there is a unique map from $\langle \mathbf{I}, \bot, \top, ^\wedge, ^\vee \rangle$ to $\langle \mathbf{I}, \top, \bot, ^\vee, ^\wedge \rangle$ that effects those interchanges and it is necessarily an involution. It is the symmetry being sought.

The $\leq$ relation on I may be defined as the most inclusive binary relation preserved by $^\wedge$ and $^\vee$ that avoids $\top \leq \bot$. We will delay the (more difficult) proof that the characterization yields a construction of the midpoint operator that figures so prominently in the opening (and throughout this work).

The assertion that the final coalgebra may be taken as the unit interval [11] needs a full proof—actually several proofs depending on the extent of constructive meaning one desires in his notion of the closed interval (see Sections 29–31, p62–75). But we move now from the coalgebraic theory with its disjunctive condition to an algebraic theory in the usual purely equational sense.

2. The Equational Theory of Scales

The theory of **scales** is given by:

a constant **top**
denoted $\top$;
a unary operation **dotting**
whose values are denoted $\dot{x}$;
a unary operation **top-zooming**
whose values are denoted $\hat{x}$ and;
a binary operation **midpointing**
whose values are denoted $x|y$

Define:

the constant **bottom**
by $\bot = \dot{\top}$;
the constant **center**
by $\odot = \bot|\top$ and;
the unary operation **bottom-zooming**
by $\overset{\vee}{x} = \dot{\overset{\wedge}{\dot{x}}}$.

We'll use both $|$ and $\mid$ for midpointing and—when convenient—denote $(x|y)^\wedge$ as $x \overset{\wedge}{|} y$ or $\widehat{x\,|\,y}$ and $(x|y)^\vee$ as $x \overset{\vee}{|} y$. Indeed, we will often treat $\overset{\wedge}{|}$ and $\overset{\vee}{|}$ as binary operations.[12]

---

[11] I is also the final coalgebra of any finite iteration of ordered wedges. If we take the n-fold ordered wedge, $X \vee X \vee \cdots \vee X$, as the set of n-tuples of the form $\langle x_0, x_1, \ldots, x_{n-1} \rangle$ such that either $x_i = \top$ or $x_{i+1} = \bot$ for $i = 0, 1, \ldots, n-2$ then the coalgebra structure is an n-tuple of functions $z_0, z_1, \ldots, z_{n-1}$ such that for each $x \in X$ and each $i = 0, 1, \ldots, n-2$ either $z_i(x) = \top$ or $z_{i+1}(x) = \bot$. The coalgebra structure on the unit interval is given by $z_i(x) = \max\big(0, \min(nx-i, 1)\big)$. Given $x \in X$ obtain the base-n expansion for its corresponding element in I by iterating (forever) the following procedure:

```
Let i = 0;
While z_i(x) = ⊤ and i < n−1 replace i with i+1;
Emit "i";
Replace x with z_i(x).
```

[12] Indeed—in my naive innocence, much to my surprise—they are associative binary operations. See [44], p17.

The six axioms (7 equations) that define the **theory of scales**:

IDEMPOTENT:

$$x|x = x$$

COMMUTATIVE: [13]

$$x|y = y|x$$

MEDIAL: [14][15]

$$(v|w) \,\Big|\, (x|y) = (v|x) \,\Big|\, (w|y)$$

CONSTANT: [16]

$$\dot{x}\;|x = \odot$$

UNITAL: [17]

$$\top \overset{\wedge}{|}\, x = x = \bot \overset{\vee}{|}\, x$$

THE SCALE IDENTITY:

$$\overset{\frown}{x|y} = \overset{\frown}{\overset{\vee}{x} \mid \hat{y}} \,\Big|\, \overset{\frown}{\hat{x} \mid \overset{\vee}{y}}$$

The **standard model** is the closed real interval $\mathbf{I}$ of all real numbers from -1 through $+1$. More generally, let $\mathbb{D}$ be the ring of dyadic rationals (those with denominator a power of 2). In **$\mathbb{D}$-modules** (or as we will pronounce them, "**dy-modules**") with total orderings we may choose elements $\bot < \top$, and define a scale as the set of all elements from $\bot$ through $\top$ with $x|y = (x{+}y)/2$, $\dot{x} = \bot + \top - x$, $\hat{x} = \max(2x{-}\top, \bot)$ (hence $\overset{\vee}{x} = \min(2x{-}\bot, \top)$). [18] The **standard interval** in $\mathbb{D}$, that is, the interval from -1 through $+1$, will be shown in the next section to be isomorphic to the **initial scale** (the scale freely generated by its constants). It will be denoted $\mathbb{I}$, the **standard $\mathbb{D}$-interval** (pronounced "**dy-interval**").[19]

The verification of all but the last of the defining equations on the standard interval is entirely routine. It will take a while before the scale identity reveals its secrets: how it first became known; how it can be best viewed; why it is true for the standard models.[20]

[13] This axiom can be replaced with a single instance: $\bot|\top = \top|\bot$. See [28] on p10.

[14] Sometimes "middle-two interchange."

[15] The medial law has a geometric interpretation: it says that the midpoints of a cycle of four edges on a tetrahedron are the vertices of a parallelogram ("the medial parallelogram"). That is, view four points A, B, C, D in general position in Euclidean space of dimension three (or more). Consider the closed path from A to B to D to C back to A and note that the four successive midpoints A|B, B|D, D|C, C|A appear in the medial law (A|B) | (C|D) = (A|C) | (B|D). This equation says, among other things, that the two line segments, the one from A|B to C|D and the one from A|C to B|D, having a point in common, are coplanar, forcing the four midpoints, A|B, B|D, D|C, C|A to be coplanar. The medial law says, further, that these two coplanar line segments have their midpoints in common, And that says—indeed, is equivalent with—A|B, B|D, D|C, C|A being the vertices of a parallelogram. (A traditional proof is obtainable from the observation that the two line segments, the one from A|B to A|C and the other from D|B to D|C, are both parallel to—and half the length of—the line segment from B to C.)

[16] A technically simpler equation is the two-variable $\dot{u}|u = \dot{v}\,|v$.

[17] The commutative axiom can be removed entirely if the first (left) unital law is replaced with $\dot{\bot} \overset{\wedge}{|}\, x = x$. See Section 28 (p61–62).

[18] In Section 14 (p33–34) we will see that every scale has a faithful representation into a product of scales that arise in this way.

[19] The free scale on one generator will be shown in Section 20 (p45–47), to be isomorphic to the scale of continuous piecewise affine functions (usually called piecewise linear) from $\mathbf{I}$ to $\mathbf{I}$ where each affine piece is given using dyadic rationals. We will give a generalization of the notion of piecewise affine so that the result generalizes: a free scale on n generators is isomorphic to the scale of all functions from $\mathbf{I}^n$ to $\mathbf{I}$ that are continuous piecewise affine with each piece given by dyadic rationals. The result further generalizes: essentially for every finitely presented scale there is a finite simplicial complex such that the scale is isomorphic to the scale of continuous piecewise affine maps with dyadic coefficients from the complex to $\mathbf{I}$. Their full subcategory can then be described in a piecewise affine manner. See Sections 21–22 (p48–51).

[20] As forbidding as the scale identity appears, this writer, at least, finds comfort in the fact that the Jacobi identity for Lie algebras looks at first sight no less forbidding. Indeed, the scale identity has 2 variables and the standard Jacobi identity 3 (with each variable appearing three times in each); the scale identity has 1 binary operation and it appears 4 times, the Jacobi has 2 binary operations, one of which appears twice and the other 6 times. (It's more efficient—and meaningful form is a bit simpler: $[[x, y], z] = [x, [y, z]] - [y, [x, z]]$.) By these counts even the high-school distributivity law is worse than the scale identity (it has 3 variables that appear a total of 7 times and 2 binary operations that appear a total of five times). It is only when the unary operations are counted that the scale identity looks worse.

An *ad hoc* verification of the scale identity on the standard model may be obtained by noting first that:

$$\hat{x} = \begin{cases} -1 & \text{if } x \leqq 0 \\ 2x-1 & \text{if } x \geqq 0 \end{cases} \qquad\qquad \overset{\vee}{x} = \begin{cases} 2x+1 & \text{if } x \leqq 0 \\ +1 & \text{if } x \geqq 0 \end{cases}$$

The scale identity then separates into four cases depending on the signatures of the two variables. When both are positive the two sides of the identity quickly reduce to $x + y - 1$ and, when both are negative, to -1. In the mixed case (because of commutativity we know in advance that the two mixed cases are equivalent) where x is positive and y is negative, the left side is, of course, $\widehat{(\frac{x+y}{2})}$ and the right side reduces to $(-1)\,\Big|\,\widehat{(x+y)}$. The entire verification is now reduced to $\widehat{z/2} = (-1)\,\Big|\,\hat{z}$ which is, in turn, quickly verified by considering the two possible signatures of z. [21]

For a fixed element a we will use $\boldsymbol{a|}$ to denote the **contraction** at a, the unary operation that sends x to $a|x$.

We will use freely:

2.1 LEMMA: SELF-DISTRIBUTIVITY:

$$a\big|(x|y) \;=\; (a|x)\;\big|\;(a|y)$$

an immediate consequence of idempotence and the medial law: $a\big|(x|y) \;=\; (a|a)\big|(x|y) \;=\; (a|x)\big|(a|y)$. This is equivalent, of course, with a contraction being a midpoint homomorphism.[22]

Define $\boldsymbol{a\triangleleft}$, the **dilatation** at a, by:

$$a\triangleleft x \;=\; \widehat{(\dot{a}\,|\bot)\overset{\vee}{|}x} \qquad [23]$$

2.2 LEMMA: *Dilatation undoes contraction:*

$$a\triangleleft(a|x) \;=\; x$$

because using the medial, constant, self-distributive and both unital laws:[24] $a\triangleleft(a|x) = \widehat{(\dot{a}\,|\bot)\overset{\vee}{|}(a|x)} = \widehat{(\dot{a}\,|a)\overset{\vee}{|}(\bot|x)} \;=\; \widehat{(\bot|\top)\overset{\vee}{|}(\bot|x)} \;=\; \widehat{\bot\overset{\vee}{|}(\top|x)} \;=\; \widehat{\top|x} \;=\; x.$ [25][26]

[21] In Section 15 (p34–36) we will show that the defining equations for scales are complete, that is, any new equation involving only the operators under discussion is either a consequence of the given equations or is inconsistent with them. Put another way: any equation not a consequence of these axioms fails in every non-degenerate scale. In particular, it fails in the initial scale. Put still another way: any equation true for any non-degenerate scale is true in all scales. A consequence is that the equational theory is decidable. In Section 27 (p60–61) it will be shown that the verification of an equation, though decidable, is NP-complete.

It may be noted that the previously stated faithful representation of free scales as scales of functions makes the equational completeness clear: if an equation on n variables fails anywhere it fails in the free scale on n generators; but if the two sides of the failed equation are not equal when represented as functions from the n-cube we may apply the evaluation operator at a dyadic-rational point where the functions disagree to obtain two distinct points in the initial scale, $\mathbb{I}$. Because the evaluation operator is a homomorphism of scales we thus obtain a counterexample in $\mathbb{I}$. Hence the set of equations that hold in all scales is the same as the set of equations that hold in $\mathbb{I}$, necessarily a complete equational theory. (Alas, the proof of the faithfulness of the representation in question requires the equational completeness.)

[22] It's worth finding the high-school–geometry proof for self-distributivity in the case that a, x and y are points in $\mathbb{R}^2$

[23] For one way of finding this formula for dilatation see [45] (p17).

[24] See Section 46 (p114–119). for a subscoring of the following equations.

[25] The zooming operations may be viewed as special cases of dilatations. One can easily verify that $\top\triangleleft x = \hat{x}$ and in [33] (p11) we'll verify that $\bot\triangleleft x = \overset{\vee}{x}$.

[26] For those looking for a Mal′cev operator, $txyz$, note that $y\triangleleft(x|z)$ is exactly that $(tabb = a = tbba)$. On that subject see New Entry 2.137 at `http://www.math.upenn.edu/~pjf/amplifications.pdf`

[27] Existential problem.

We immediately obtain:

2.3 LEMMA: THE CANCELLATION LAW:

$$\textit{If}\quad a|x = a|y \quad \textit{then}\quad x = y.$$

Two important equations for dotting:

2.4 LEMMA: THE INVOLUTORY LAW:

$$\ddot{x} = x$$

2.5 LEMMA: DOT-DISTRIBUTIVITY:

$$(u\,|\,v)^{\bullet} = \dot{u}\,\dot{|}\,\dot{v}$$

Both can be quickly verified using commutativity and cancellation: $\dot{x}\,|\,\ddot{x} = \ddot{x}\,|\,\dot{x} = \dot{x}\,|x$ and $(u|v)\big|(u|v)^{\bullet} = (u|v)^{\bullet}\big|(u|v) = \dot{u}\,|u = (\dot{u}\,|u)\big|(\dot{u}\,|u) = (\dot{u}\,|u)\,\big|(\dot{v}\,|v) = (u\,|\,\dot{u})\big|(v\,|\,\dot{v}) = (u|v)\big|(\dot{u}\,|\,\dot{v})$. [28]

Given a term $txy\ldots z$ involving $\top$, $\bot$, midpointing, dotting, $\top$- and $\bot$-zooming, the **dual term** is the result of fully applying the distributivity and involutory laws to $(t\;\dot{x}\;\cdots\;\dot{z})^{\bullet}$. It has the effect of swapping $^\wedge$ with $^\vee$ and $\top$ with $\bot$. If we replace both sides of an equation with their dual terms we obtain the **dual equation**.

We have already seen one pair of dual equations, to wit, the unital laws. (That's not counting a whole bunch of self-dual equations.) The dual equation of the scale identity is:

$$u\overset{\vee}{|}v \;=\; (\hat{u}\overset{\vee}{|}\overset{\vee}{v})\,\big|\,(\overset{\vee}{u}\overset{\vee}{|}\hat{v})$$

Note that we have not yet allowed dilatations in the terms to be dualized.[29]

As a direct consequence of the idempotent and unital laws we have that $\top$ is a fixed-point for $\top$-zooming: $\overset{\wedge}{\top} = \widehat{\top|\top} = \top$. Dually, $\overset{\vee}{\bot} = \bot$. $\top$ is a also a fixed-point for $\bot$-zooming using the unital law, scale identity (for the first time), idempotent, commutative and unital laws: $\top = \widehat{\top|\top} = \widehat{\overset{\vee}{\top}|\overset{\wedge}{\top} \,\big|\, \overset{\wedge}{\top}\dot{|}\overset{\vee}{\top}} = \widehat{\overset{\vee}{\top}|\top \,\big|\, \top|\overset{\vee}{\top}} = \overset{\vee}{\top}\dot{|}\overset{\vee}{\top} = \overset{\vee}{\top}$. Dually $\overset{\wedge}{\bot} = \bot$. That is:

2.6 LEMMA: *Both* $\top$ *and* $\bot$ *are fixed-points for both* $^\wedge$ *and* $^\vee$.

[28] To see how the axiom $\bot|\top = \top|\bot$ suffices for commutativity, first note that commutativity was not used to obtain the left cancellation law ($a|x = a|y$ implies $x = y$). One consequence is that $\dot{x} = \dot{v}$ implies $x = v$ (use left cancellation on $\dot{x}|x = \dot{x}|v$). Besides being monic, dotting is epic because the second unital law, $\bot\overset{\vee}{|}\,x = x$, when written in full says $\Big(\big((\bot|x)^{\bullet}\big)^{\wedge}\Big)^{\bullet} = x$, hence for all x there is v such that $\dot{v} = x$ (to wit, $\big((\bot|x)^{\bullet}\big)^{\wedge}$). Hence dotting is an invertible operation.

If $y|x = \odot$ then $y = \dot{x}$ because if we let z be such that $\dot{z} = y$ then $y|x = y|z$ and we use cancellation to obtain $x = z$ and, hence, $y = \dot{x}$. A consequence is $(u|v)^{\bullet} = \dot{u}\,|\,\dot{v}$ because it suffices to show $(\dot{u}\,|\,\dot{v})\big|(u|v) = \odot$ which follows easily using the medial, constant and idempotence laws.

The commutativity of $\top$ and $\bot$ says $\top|\bot = \odot$ hence $\top = \dot{\bot}$ and that equation when combined with dot-distributivity and the second unital law yields $x = \bot\overset{\vee}{|}\,x = \Big(\big((\bot|x)^{\bullet}\big)^{\wedge}\Big)^{\bullet} = \Big(\big(\dot{\bot}|\dot{x}\big)^{\wedge}\Big)^{\bullet} = \Big(\big(\top|\dot{x}\big)^{\wedge}\Big)^{\bullet} = \ddot{x}$, quite enough to establish that the center is central: $x|\odot = (x|x)\big|(\dot{x}|x) = (x|\dot{x})\big|(x|x) = (\ddot{x}\,|\,\dot{x})\big|x = \odot|x$. Finally, $\odot\big|(x|y) = (\odot|x)\big|(\odot|y) = (\odot|x)\big|(y|\odot) = (\odot|y)\big|(x|\odot) = (\odot|y)\big|(\odot|x) = \odot\big|(y|x)$ and left cancellation yields $x|y = y|x$. See Section 46 (p114–119) for subscorings.

[29] In time we will be able to do so. We will show (in [56], p23) that dilatations are self-dual just as are midpointing and dotting. That is, we will show $\Big(\big((\dot{a}\,|\bot)|x\big)^{\vee}\Big)^{\wedge} = \Big(\big((\dot{a}\,|\top)|x\big)^{\wedge}\Big)^{\vee}$.

In our second direct use of the scale identity we replace its second variable with $\top$ and use the unital law to obtain:

2.7 LEMMA: THE LAW OF COMPENSATION:

$$x \;=\; \overset{\vee}{x}\,\big|\,\hat{x}$$

because $x = \overset{\frown}{x|\top} = \overset{\frown}{\overset{\vee}{x}|\overset{\wedge}{\top}} \,\big|\, \overset{\frown}{\hat{x}|\overset{\vee}{\top}} = \overset{\frown}{\overset{\vee}{x}|\top} \,\big|\, \overset{\frown}{\hat{x}|\top} = \overset{\vee}{x}\,|\,\hat{x}$.

A consequence:

2.8 LEMMA: THE ABSORBING LAWS:

$$x \overset{\wedge}{|} \bot \;=\; \bot \qquad \textit{and} \qquad x \overset{\vee}{|} \top \;=\; \top$$

because we can use the law of compensation and then cancellation on $x|\bot = (x \overset{\vee}{|} \bot) \,\big|\, (x \overset{\wedge}{|} \bot) = x\big|(x \overset{\wedge}{|} \bot)$. [31]

2.9 LEMMA: *A scale is trivial iff* $\top = \bot$.

Because if $\top = \bot$ then $x = \overset{\frown}{x|\top} = \overset{\frown}{x|\bot} = \bot$ for all x.

The unital laws (or, for that matter, the absorbing laws) easily yield:

2.10 LEMMA:

$$\overset{\wedge}{\odot} \;=\; \bot \qquad \textit{and} \qquad \overset{\vee}{\odot} \;=\; \top$$

The center is the only self-dual element, $\dot{\odot} = \odot$. (If $\dot{x} = x$ then apply $x|$ to both sides to obtain $x|\dot{x} = x|x$, that is, $\odot = x$.)

If the second variable is replaced with $\odot$ in the scale identity (this is its third direct use [32]), the equations $\overset{\wedge}{\odot} = \bot$ and $\overset{\vee}{\odot} = \top$ yield a special case of the (not correct-in-general) distributive laws for $\top$- and $\bot$-zooming:

2.11 LEMMA: THE CENTRAL DISTRIBUTIVITY LAWS:

$$x \overset{\wedge}{|} \odot \;=\; \hat{x}\,|\bot \qquad \textit{and} \qquad x \overset{\vee}{|} \odot \;=\; \overset{\vee}{x}\,|\top$$

because $\overset{\frown}{x|\odot} = \overset{\frown}{\overset{\vee}{x}|\overset{\wedge}{\odot}} \,\big|\, \overset{\frown}{\hat{x}|\overset{\vee}{\odot}} = \overset{\frown}{\overset{\vee}{x}|\bot} \,\big|\, \overset{\frown}{\hat{x}|\top} = \bot|\hat{x}$. [33]

We will need:

2.12 LEMMA:

$$x \;=\; y \qquad \textit{iff} \qquad \dot{x}\,|\,y \;=\; \odot$$

because if $\dot{x}\,|\,y = \odot$ we can use cancellation on $\dot{x}\,|\,y = \dot{x}\,|\,x$.

A consequence is what's called "**swap-and-dot**:" given $w|x = v|z$ swap-and-dot any pair of terms from opposite sides to obtain equations such as $w|\dot{v} = \dot{x}|z$. (From $w|x = v|z$ infer $\odot = (w|x)\big|(v|z)^{\cdot} = (w|x)\big|(\dot{v}\,|\,\dot{z}) = (w|\dot{v})\big|(x|\dot{z}) = (w|\dot{v})\big|(\dot{x}\,|z)^{\cdot}$.)

Note that the commutative and medial laws say that $(w|x)\big|(y|z)$ is invariant under all 24 permutations of the variables (as, of course, are $(w|x)\overset{\vee}{|}(y|z)$ and $(w|x)\overset{\wedge}{|}(y|z)$).

[31] $\overset{\vee}{\top} = \top$ can now be viewed as a special case of the absorbing law: $\overset{\vee}{\top} = \top \overset{\vee}{|} \top = \top$.

[32] This is the last direct use of the scale identity until Section 4 (p14–17).

[33] As promised, we can now easily prove $\bot \lhd x = \big((\dot{\bot}\,|\bot)|x\big)^{\vee\wedge} = \big(\odot|x\big)^{\vee\wedge} = \big(\top|\overset{\vee}{x}\big)^{\wedge} = \overset{\vee}{x}$.

3. The Initial Scale

3.1 THEOREM: *The standard* $\mathbb{D}$*-interval* $\mathbb{I}$ *(the standard interval of dyadic rationals) is isomorphic to the initial scale and it is simple.*

(Recall that for any equational theory "simple" means having exactly two quotient structures, the identity and the trivial.) When coupled with the previous observation that $\perp \neq \top$ in all non-trivial scales we thus obtain:

3.2 THEOREM: $\mathbb{I}$ *appears uniquely as a subscale of every non-trivial scale.*

The proof is on the computational side as, apparently, it must be. It turns out that not all of the axioms are needed for the proof and that leads to another theorem of interest in which—for technical reasons—we define the theory of **minor scales** to be the result of removing the scale identity but adding the absorbing laws (either one, by itself, would suffice).[34]

3.3 THEOREM: *The theory of minor scales has a unique equational completion (and using Theorem* 15.1 (p34)*, that complete theory is the theory of scales).*

There are several ways of restating this fact: equations consistent with the theory are consistent with each other; an equation is true for all scales iff it holds for any non-trivial minor scale; an equation is true for all scales iff it is consistent with the theory of minor scales; using the completeness of the theory of scales, every equation is either inconsistent with the theory of minor scales or is a consequence of the scale identity.[35]

The proof is obtained by showing that the initial minor scale is $\mathbb{I}$ and is simple. (Thus every consistent extension of the theory of minor scales, having a non-degenerate model, must hold for every subalgebra, hence must hold for the initial model. The complete equational theory of the initial model thus includes all equations consistent with the theory of minor scales.)[36]

---

[34] "Technical reasons" means other than the existence of interesting examples. For an example of a minor scale that is not a scale see Section 28 (p61–62).

[35] The same relationship holds between the theory of lattices and the theory of distributive lattices, and between the theories of Heyting and Boolean algebras. A less well-known example: for any prime p, the unique equationally consistent extension of the theory of characteristic-p unital rings is the theory of characteristic-p unital rings satisfying the further equation $x^p = x$. (This almost remains true when the unit is dropped: given a maximal consistent extension of the theory of "rngs" there is a prime p such that the theory is either the theory of characteristic-p rngs that satisfy the same equation as above ($x^p = x$), or the theory of elementary p-groups with trivial multiplication ($xy = 0$).)

A telling pair of examples: the equational theory of lattice-ordered groups and the equational theory of lattice-ordered rings. In each case the unique maximal consistent equational extension is the set of equations that hold for the integers. The first case is decidable (and all one needs to add to obtain a complete set of axioms is the commutativity of the group operation—see Section 28, p61–62). The second case is undecidable: the non-solvability of any Diophantine equation, $P = 0$, is equivalent to the consistency of the equation $1 \wedge P^2 = 1$. (Conversely, one may show that the consistency of any equation is equivalent to the non-solvability of some Diophantine equation, for instance, by showing that the solvability of an equation involving lattice operations is always equivalent to the solvability of an equation with one fewer lattice operation: given a term S in the theory of lattice-ordered commutative rings replace S with S^2 if necessary, to insure that it has no negative values and let $P \diamond Q$ be an inner-most instance of a lattice operation—that is, one in which P and Q are ordinary polynomials; let T be the result of replacing $P \diamond Q$ in S with a fresh variable z; let a, b, c, d be four more fresh variables; let e denote the value of $1 \diamond (-1)$; then the solvability of $S = 0$ in $\mathbb{Z}$ is equivalent to the solvability of $T + \left[P + Q + e(a^2 + b^2 + c^2 + d^2) - 2z\right]^2 + \left[(a^2 + b^2 + c^2 + d^2)^2 - (P - Q)^2\right]^2 = 0$.)

[36] We can not only drop axioms but structure: $\mathbb{I}$ is the free **midpoint algebra** on two generators ($\top$,$\perp$) and the free **symmetric midpoint algebra** on one generator where we understand the first three scale equations (idempotent, commutative and medial) to define midpoint algebras and the first four (add the constant law) together with the involutory and distributive laws for dotting to define symmetric midpoint algebras. In the opening section I talked about **closed midpoint algebras** with reference to the structure embodied by top, bottom and midpointing, with the remark that the axioms were not needed in the material of that section. Let me now legislate that the axioms are the first three scale equations together with the non-equational Horn sentence of cancellation for midpointing. Since $\mathbb{I}$ is such, it will perforce be the case that $\mathbb{I}$ is the initial closed midpoint algebra. (The set $\{\perp, \top\}$ is a two-element midpoint algebra but not a closed midpoint algebra when we take $\perp|\top = \perp$ and it is a symmetric midpoint algebra when we take $\dot{\top} = \perp$.) For a **symmetric closed midpoint algebra** add dotting and the constant law (the involutory and distributive laws are consequences of cancellation). $\mathbb{I}$ is also the initial symmetric closed midpoint algebra.

It should be noted, however, that there are closed midpoint algebras, even symmetric closed midpoint algebras, that extend the notion of midpoint. Choose an odd number of evenly spaced points on a circle and define the midpoint of any two of them to be the unique equidistant point in the collection. Chose any two points for $\top$ and $\perp$ and define $\dot{x}$ to be the unique element such that $\dot{x}\,|x = \perp|\top$. If one chooses $\perp, \perp|\top, \top$ to be adjacent then the induced map from $\mathbb{I}$ is guaranteed to be

We first construct the initial minor scale via a "canonical form" theorem and show that it is simple. Define a term in the signature of scales to be of "grade -1" if it is either $\top$ or $\bot$, of "grade 0" if it is $\odot$, and of "grade $n{+}1$" if it is either $\top|A$ or $\bot|A$ where A is of grade n. We need to show that the elements named by graded terms form a subscale. Closure under the unary operations—dotting, $\top$- and $\bot$-zooming—is straightforward (but note that the absorbing laws are needed). For closure under midpointing we need an inductive proof. We consider $A|B$ where A is of grade a and B is of grade b. Because of commutativity we may assume that $a \leq b$. The induction is first on a. The case $a = -1$ presents no difficulties. For $a = 0$ we must consider two sub-cases, to wit when $b = 0$ and when $b > 0$. If $b = 0$ then $A|B = \odot$. When $b > 0$ we may, without loss of generality, assume that $B = \top|B'$ for B' of grade $b{-}1$. But then $A|B = \odot\big|(\top|B') = (\top|\bot) \bigm| (\top|B') = \top\big|(\bot|B')$ which is of grade $b{+}1$. If $a > 0$ there are, officially, four sub-cases to consider, but, without loss of generality, we may assume that $A = \top|A'$ and either $B = \top|B'$ or $B = \bot|B'$ where A' is of grade $a{-}1$ and B' is of grade $b{-}1$. In the homogeneous sub-case we have that $A|B = (\top|A') \bigm| (\top|B') = \top|(A'|B')$ and by inductive hypothesis we know that $A'|B'$ is named by a graded term, hence so is $A|B$. In the heterogeneous sub-case we have that $A|B = (\top|A') \bigm| (\bot|B') = (\top|\bot) \bigm| (A'|B') = \odot\big|(A'|B')$ and by inductive hypothesis we know that $A'|B'$ is named by a graded term and we then finish by invoking again the case $a = 0$.

The simplicity of the initial scale—and the uniqueness of graded terms—also requires induction. Suppose that A and B are distinct graded terms and that $\equiv$ is a congruence such that $A \equiv B$. Again we may assume that $a \leq b$. In the case $a = -1$ we may assume without loss of generality that $A = \top$. The sub-case $b = -1$ is, of course, the prototypical case ($B = \bot$ else $A = B$). For $b = 0$ we infer from $\top \equiv \odot$ that $\overset{\wedge}{\top} \equiv \overset{\wedge}{\odot}$, hence $\top \equiv \bot$, returning to the sub-case $b = 0$. For $b > 0$ we must consider the two sub-cases, $B = \top|B'$ and $B = \bot|B'$ where B' is of grade $b{-}1$. From $\top \equiv \top|B'$ we may infer $\overset{\wedge}{\top} \equiv \widehat{\top|B'}$, hence $\top \equiv B'$ and thus reduce to the earlier sub-case $b{-}1$. From $\top \equiv \bot|B'$ we may infer $\overset{\wedge}{\top} \equiv \widehat{\bot|B'}$ immediately reducing to the sub-case $b = -1$. For the case $a = 0$ we know from $a \leq b$ and $A \neq B$ that $b > 0$ and we may assume without loss of generality that $B = \top|B'$ where B' is of grade $b{-}1$. But $\odot \equiv B$ then says that $\overset{\wedge}{\odot} \equiv \hat{B}$, hence $\top \equiv B'$ and we reduce to the case $a = -1$. For the case $a > 0$ we again come down to two sub-cases. In the homogeneous sub-case $A = \top|A'$ and $B = \top|B'$ we infer from $A \equiv B$ that $\widehat{\top|A'} \equiv \widehat{\top|B'}$ hence that $A' \equiv B'$. Since $A \neq B$ we have that $A' \neq B'$ and we reduce to the case $a{-}1$. Finally, in the heterogeneous sub-case $A = \top|A'$ and $B = \bot|B'$ we infer from $A \equiv B$ that $\top \overset{\vee}{|} A' \equiv \bot \overset{\vee}{|} B'$ hence that $\top \equiv B'$. Since the grade of B' is positive such reduces to the case $a = -1$.

When we know that every non-trivial scale contains a minimal scale isomorphic to the initial scale, then perforce we know that there is, up to isomorphism, only one non-trivial minimal scale. Hence, to see that the initial scale is isomorphic to $\mathbb{I}$ it suffices to show that $\mathbb{I}$ is without proper subscales, or—to put it more constructively—that every element in $\mathbb{I}$ can be accounted for starting with $\top$. Switching to $\mathbb{D}$-notation, we have $1 = \top, -1 = \dot{\top}$ and $0 = \top|\bot$. We know that all other elements are of the form $\pm(2n{+}1)2^{-(m+1)}$ where n and m are natural numbers and $2n{+}1 < 2^{m+1}$ Inductively (on m):

$$\pm(2n{+}1)2^{-(m+1)} \quad = \quad \pm n2^{-m} \bigm| \pm(n{+}1)2^{-m} \qquad [37]$$

onto. $\mathbb{I}$ thus has an infinite number of closed midpoint quotients and it is far from simple (the simple algebras are precisely the cyclic examples of prime order).

[37] Using, of course, that $n{+}1 \leq 2^m$ (because, necessarily, $2n{+}2 \leq 2^{m+1}$).

4. Lattice Structure

The most primitive way of defining the natural partial order on a scale is to define $u \leq v$ iff there is an element w such that $u|\top = v|w$. From this definition it is clear that any map that preserves midpointing and $\top$ must preserve order (which together with von Neumann is quite enough to prove the opening assertion of this work: see Section 40, p91–94).[38]

But in the presence of zooming we may remove the existential. First note that

$$\exists_w \; z = \top|w \quad \text{iff} \quad \check{z} = \top$$

because if $z = \top|w$ then the absorbing law says $\check{z} = \top \mathbin{\overset{\vee}{|}} w = \top$. Conversely, if $\check{z} = \top$ then we may take $w = \hat{z}$ (the law of compensation gives us $z = \check{z} \mid \hat{z} = \top| \hat{z}$).

If we use swap-and-dot and the involutory law to rewrite the existential condition for $u \leq v$ as $\exists_w \; [\dot{u}\,|v = \top|w]$ we are led to define a new binary operation to be denoted by borrowing from J-Y Girard,[39] $\boldsymbol{u \multimap v} = \dot{u} \mathbin{\overset{\vee}{|}} v$ and we see by the absorbing law that

$$u \leq v \quad \text{iff} \quad u \multimap v \;=\; \top.$$

We make this our official definition ($u \multimap v$ may be read as "the extent to which u is less than v" where $\top$ is taken as "true" [40]). A neat way to encapsulate this material is with:

4.1 LEMMA: THE LAW OF BALANCE:

$$u \mid (u \multimap v) \;=\; v \mid (v \multimap u)$$

(One can see at once that $u \multimap v = \top$ implies that $u|\top = v|w$ is solvable.) To prove the law of balance note that the law of compensation yields $u| \dot{v} = (u \mathbin{\overset{\wedge}{|}} \dot{v})|(u \mathbin{\overset{\vee}{|}} \dot{v})$ and a swap-and-dot yields $u|\dot{(u \mathbin{\overset{\wedge}{|}} \dot{v})} = v|(u \mathbin{\overset{\vee}{|}} \dot{v})$; the left side rewrites as $u|(\dot{u} \mathbin{\overset{\vee}{|}} v) = u|(u \multimap v)$ and the right as $v|(\dot{v} \mathbin{\overset{\vee}{|}} u) = v|(v \multimap u)$.

We verify that $\leq$ is a partial order as follows:

Reflexivity is immediate: $x \multimap x = \dot{x} \mathbin{\overset{\vee}{|}} x = \check{\odot} = \top$.

For antisymmetry, given $x \multimap y = \top = y \multimap x$ just apply the unital law to both sides of the law of balance.

Transitivity is not so immediate. We will need that $\check{u} = \top = \check{v}$ implies $u \mathbin{\overset{\vee}{|}} v = \top$ (true because, using the law of compensation, $\check{u} = \top = \check{v}$ says $u \mathbin{\overset{\vee}{|}} v = (\check{u}|\hat{u}) \mathbin{\overset{\vee}{|}} (\check{v}|\hat{v}) = (\top|\hat{u}) \mathbin{\overset{\vee}{|}} (\top|\hat{v}) = \top \mathbin{\overset{\vee}{|}} (\hat{u}|\hat{v}) = \top$). Hence if $u \multimap v$ and $v \multimap w$ are both $\top$ then so is $(\dot{u}|v) \mathbin{\overset{\vee}{|}} (\dot{v}|w)$. But this last term is equal (using the commutative, medial and constant laws) to $\odot \mathbin{\overset{\vee}{|}} (\dot{u}\,|w)$ which by the central distributivity law is $\top|(\dot{u} \mathbin{\overset{\vee}{|}} w)$. Hence if $u \leq v$ and $v \leq w$ we have that $\top|(u \multimap w) = \top$ and when both sides are $\top$-zoomed we obtain $u \multimap w = \top$.

Covariance of $z|$ also follows from central distributivity:

$$(z|x) \multimap (z|y) = (z|x)^{\cdot} \mathbin{\overset{\vee}{|}} (z|y) = (\dot{z}\,|\,\dot{x}) \mathbin{\overset{\vee}{|}} (z|y) = (\dot{z}\,|z) \mathbin{\overset{\vee}{|}} (\dot{x}\,|y) = \odot \mathbin{\overset{\vee}{|}} (\dot{x}\,|y) = \top|(x \multimap y) \quad [41]$$

Hence $x \leq y$ implies $z|x \leq z|y$. Not only does $z|$ preserve order, it also reflects it.

[38] Left as an easy exercise: a map that preserves midpointing and $\perp$ also preserves order.
[39] For reasons to become clear in the next section.
[40] Again, see the next section.
[41] See Section 46 (p114–119) for a subscoring.

Contravariance of dotting is immediate:

$$\dot{u} \multimap \dot{v} \;=\; v \multimap u.$$

For each semiquation we obtain the **dual semiquation** by replacing the terms with their duals and reversing the semiquation.

A few more formulas worth noting are:

$$\begin{aligned} \top \multimap x &= x \\ x \multimap \bot &= \dot{x} \\ \dot{x} \multimap x &= \check{x} \\ x \multimap \dot{x} &= \dot{\hat{x}} \end{aligned}$$

Note that $\odot \multimap x = \top|\check{x}$ hence $\check{x} = \dot{\top}$ iff $\odot \leq x$. The lemma we needed (and proved) for transitivity, that $\check{u} = \top = \check{v}$ implies $u \curlyvee v = \top$, is now an easy consequence of the covariance of midpointing.

It is immediate from the definition and absorbing laws that $\bot \leq x \leq \top$ all x.

We obtain a swap-and-dot lemma for semiquations:

4.2 LEMMA:

$$u|v \;\leq\; w|x \quad \textit{iff} \quad u|\dot{w} \;\leq\; \dot{v}|x$$

Because $(u|v) \multimap (w|x) \;=\; (u|v)^{\bullet} \curlyvee (w|x) \;=\; (\dot{u}|\dot{v}) \curlyvee (w|x) \;=\; (\dot{u}|w) \curlyvee (\dot{v}|x) \;=\; (u|\dot{w})^{\bullet} \curlyvee (\dot{v}|x) \;=\; (u|\dot{w}) \multimap (\dot{v}|x)$. (See Section 46, p114–119 for a subscoring.)

An important fact: $\top$ is an **extreme point** in the convex-set sense, that is, it is not the midpoint of other points (see Section 41, p95, for a discussion on this definition):

4.3 LEMMA: $x|y = \top$ *iff* $x = \top = y$.

Because $x|y \leq \top|y$ (without any hypothesis), hence $x|y = \top$ implies $\top = x|y \leq \top|y \leq \top$ forcing $\top|y = \top$, hence $y = \widehat{\top|y} = \hat{\top} = \top$.

4.4 LEMMA: *Zooming is covariant.*

The covariance of $\top$-zooming (and, hence, $\bot$-zooming) requires work. In constructing this theory an equational condition was needed that would yield the Horn condition that $u \leq v$ implies $\hat{u} \leq \hat{v}$. (The equation, for example, $(u \multimap v) \multimap (\hat{u} \multimap \hat{v}) = \top$ would certainly suffice. Alas, this equation is inconsistent with even the axioms of minor scales: if we replace u with $\top$ and v with $\odot$ it becomes the assertion that $\odot \leq \bot$.)

The fact that $\top$ is an extreme point says that it would suffice to have:

$$u \multimap v \;=\; (\hat{u} \multimap \hat{v}) \,\big|\, (\check{u} \multimap \check{v}).$$

Indeed, this equation implies that the two zooming operations collectively preserve and *reflect* the order.

Finding a condition strong enough is, as noted, easy. To check that it is not *too* strong, that is to check the equation on the standard model, it helps to translate back to more primitive terms, in which we wish to prove that these are equivalent

$$\dot{u} \curlyvee v \;=\; (\dot{\hat{u}} \curlyvee \hat{v}) \,\big|\, (\dot{\check{u}} \curlyvee \check{v}) \;=\; (\check{\dot{u}} \curlyvee \hat{v}) \,\big|\, (\hat{\dot{u}} \curlyvee \check{v}).$$

Since dotting is involutory such is equivalent to the last equation but without the dots:

$$u \overset{\vee}{|} v \;=\; (\overset{\vee}{u} \overset{\vee}{|} \hat{v}) \;\big|\; (\hat{u} \overset{\vee}{|} \overset{\vee}{v})$$

to wit, the dual of the scale identity. And it was this that was the first appearance of the scale identity (and its first serious use—the three previous direct applications that have appeared here served only to replace what had, in fact, once been axioms, to wit, the variable-free equation $\overset{\vee}{\bot} = \bot$ and the two one-variable equations, $\overset{\vee}{x} \mid \hat{x} = x$ and $\widehat{\odot|x} = \bot| \hat{x}$, which three laws are much more apparent than is the scale identity).

Among the corollaries are the covariance of the binary operations $\overset{\wedge}{|}$, $\overset{\vee}{|}$ and the important semiquations:

4.5 LEMMA:

$$\hat{x} \;\leq\; x \;\leq\; \overset{\vee}{x}$$

Because $\hat{x} = \; x \overset{\wedge}{|} x \;\leq\; \top \overset{\wedge}{|} x \;=\; x \;=\; x \overset{\vee}{|} \bot \;\leq\; x \overset{\vee}{|} x \;= \overset{\vee}{x}$.

Further corollaries: $x \multimap y$ is covariant in y and contravariant in x. (If $a \triangleleft x$ is viewed as a binary operation then it is covariant in x and contravariant in a.)

4.6 LEMMA: THE CONVEXITY OF $\top$-ZOOMING:

$$\widehat{u\,|\,v} \;\leq\; \hat{u}\,|\,\hat{v}$$

Because $u \overset{\wedge}{|} v \;=\; (\overset{\vee}{u}|\hat{u}) \overset{\wedge}{|} (\overset{\vee}{v}|\hat{v}) \;\leq\; (\top|\hat{u}) \overset{\wedge}{|} (\top|\hat{v}) \;=\; \top \overset{\wedge}{|} (\hat{u}\,|\,\hat{v}) \;=\; \hat{u}\,|\,\hat{v}$.

The dual semiquation:

4.7 LEMMA:

$$u \overset{\vee}{|} v \;\geq\; \overset{\vee}{u}\,|\,\overset{\vee}{v}$$

Define a binary operation, temporarily denoted $x \diamond y$, as $x \overset{\wedge}{|} (x \multimap y)$. The law of balance says, in particular, that the $\diamond$-operation is commutative and consequently covariant not just in y but in both variables. Note that if $x \leq y$ then $x \diamond y \,=\, x \overset{\wedge}{|} (x \multimap y) = x \overset{\wedge}{|} \top = x$, which together with commutativity says that whenever x and y are comparable, $x \diamond y$ is the smaller of the two. As special cases we obtain the three equations: $x \diamond \top \;=\; \top \diamond x \;=\; x \diamond x \;=\; x$. These three equations together with the covariance are, in turn, enough to imply that $x \diamond y$ is the greatest lower bound of x and y: from the covariance and $x \diamond \top = x$ we may infer that $x \diamond y \leq x \diamond \top = x$ and, similarly, $x \diamond y \leq y$; from covariance and $z \diamond z = z$ we may infer that $x \diamond y$ is the greatest lower bound (because $z \leq x$ plus $z \leq y$ implies $z = z \diamond z \leq x \diamond y$).

All of which gives us the lattice operations (using duality for $x \vee y$):

4.8 LEMMA:

$$x \wedge y \;=\; x \overset{\wedge}{|} (x \multimap y) \;=\; x \overset{\wedge}{|} (\dot{x} \overset{\vee}{|} y) \quad [42]$$

$$x \vee y \;=\; x \overset{\vee}{|} \widehat{\dot{x}|y}$$

We will extend the notion of duality to include the lattice structure. (But note that we do not have a symbol for the dual of $\multimap$.)

Direct computation now yields what we will see must be known by an oxymoronic name; it is the "internalization" of the (external) disjunction: $\overset{\vee}{x} = \top$ *or* $\hat{x} = \bot$.

[42] It behooves us to figure out just what the term $x\big|(x \multimap y)$ is before it is $\top$-zoomed. We know that it is commutative and covariant in both variables. The law of compensation says that it is equal to $(x \overset{\vee}{|} (x \multimap y)) \;\big|\; (x \overset{\wedge}{|} (x \multimap y))$. We know now what the right-hand term, $x \overset{\wedge}{|} (x \multimap y)$, is. For the left-hand term, $x \overset{\vee}{|} (x \multimap y)$, note that its covariance implies that it is always at least $\bot \overset{\vee}{|} (\bot \multimap \bot) = \bot \overset{\vee}{|} \top = \top$. Hence $x\big|(x \multimap y) = \top\big|(x \wedge y)$.

4.9 LEMMA: THE COALGEBRA EQUATION:

$$\overset{\vee}{x} \vee \dot{\hat{x}} = \top$$

because: $\overset{\vee}{x} \vee \dot{\hat{x}} = \overset{\vee}{x} \overset{\vee}{|} \widehat{\dot{\overset{\vee}{x}} \,|\, \dot{\hat{x}}} = \overset{\vee}{x} \overset{\vee}{|} \widehat{(\overset{\vee}{x}|\hat{x})^{\bullet}} = \overset{\vee}{x} \overset{\vee}{|} \widehat{\dot{x}} = (\overset{\vee}{x}|\dot{\overset{\vee}{x}})^{\vee} = \overset{\vee}{\odot} = \top.$

If we replace x with $\dot{x}|y$ we obtain the internalization of the disjunction: $(x \leq y)$ *or* $(y \leq x)$.

4.10 LEMMA: THE EQUATION OF LINEARITY:

$$(x \multimap y) \vee (y \multimap x) = \top$$ [43]

Indeed it says that what logicians call the **disjunction property** (the principle that a disjunction equals $\top$ only if one of the terms equals $\top$) is equivalent with linearity:

4.11 LEMMA: *The following are equivalent for scales:*
Linearity; The disjunction property; The coalgebra condition.

We will need:

4.12 LEMMA: THE ADJOINTNESS LEMMA:

$$u \leq v \multimap w \quad \textit{iff} \quad u \overset{\wedge}{|} v \leq w$$

Because if $u \leq v \multimap w$ then $v \overset{\wedge}{|} u \leq v \overset{\wedge}{|} (v \multimap w) = v \wedge w \leq w$. And if $u \overset{\wedge}{|} v \leq w$ then $u \leq \dot{v} \vee u = \dot{v} \overset{\vee}{|} \widehat{v|u} = v \multimap \widehat{u|v} \leq v \multimap w$. [44]

We close this section with a few **interval isomorphisms**. For any $b < t$ the interval $[b, t|$ is order-isomorphic with an interval whose top end-point is $\top$, to wit, the interval $[t \multimap b, \top]$. The isomorphism is $t \multimap (-)$. Its inverse is $\dot{t} \overset{\wedge}{|} (-)$. (The fact that the composition $\dot{t} \overset{\wedge}{|} (t \multimap x) = x$ for all $x \in [b, t]$ is just the fact that $t \overset{\wedge}{|} (t \multimap x) = t \wedge x = x$. The fact that the composition $t \multimap (t \overset{\wedge}{|} x) = x$ for all $x \in [t \multimap b, \top]$ is just the fact that $\dot{t} = t \multimap \bot \leq t \multimap b \leq x$ and hence that $t \multimap (t \overset{\wedge}{|} x) = \dot{t} \overset{\vee}{|} (t \overset{\wedge}{|} x) = \dot{t} \vee x = x$. These are not just order-isomorphisms. With the forthcoming Theorem 15.1 (p34) it will be easy to prove that they preserve midpointing and when—in Section 17 (p41)—we note that all closed intervals have intrinsic scale structures it will be easy to see that they are scale isomorphisms.[45]

[43] The coalgebra equation is obtainable, in turn, from the equation of linearity by replacing y with $\dot{x}$.

[44] Using the linear representation theorem from Section 8 (p22–25) one can show the initially surprising fact that the binary operation $\overset{\wedge}{|}$ is associative. Any poset may be viewed as a category and this associativity together with the adjointness lemma allows us to view a scale as a "symmetric monoidal closed category" with $\overset{\wedge}{|}$ as the monoidal product and $\multimap$ as the "closed" structure. The monoidal unit is $\top$. A scale is, in fact, a "$\star$-autonomous category": the "dualizing object" is $\bot$.

A straightforward verification of the associativity of $\overset{\wedge}{|}$ on a linear scale entails a lot of case analysis. Perhaps it is best to use the equational completion that will be proved in Section 15 (p34–36); it suffices to verify it on just one non-trivial example. The easiest we have found is to take the unit interval—not the standard interval—and to verify the dual equation, the associativity of $\overset{\vee}{|}$. On the unit interval $x \overset{\vee}{|} y$ is addition truncated at 1, easily seen to be associative.

With the associativity in hand there's a nicer proof for the adjointness, even better, for the internal version of the adjointness:
$u \multimap (v \multimap w) = \dot{u} \overset{\vee}{|} (\dot{v} \overset{\vee}{|} w) = (\dot{u} \overset{\vee}{|} \dot{v}) \overset{\vee}{|} w = (u \overset{\wedge}{|} v)^{\bullet} \overset{\vee}{|} w = (u \overset{\wedge}{|} v) \multimap w.$

[45] The construction of the dilatation operator can be motivated by this material. For any a, the function $(a|\top) \multimap (-)$ sends the image of $a|-$ to the interval $[\odot, \top]$, quite enough to suggest that $(a|\top) \multimap (a|-)$ is the same as $\top|-$, hence (using the unital law) that $\widehat{(a|\top) \multimap (a|x)} = x$. The function $\widehat{(a|\top) \multimap (-)}$ is $a \triangleleft -$.

The construction for dilatation came rather late for me, but not as late as interval isomorphisms. It arose from the observation that each of the "quarter-intervals" on $[-1, 1]$, that is, the intervals $[-1, -\frac{1}{2}]$, $[-\frac{1}{2}, 0]$, $[0, \frac{1}{2}]$, $[\frac{1}{2}, 1]$, are isomorphic to the entire interval via a pair of zoomings, hence if we could move the image of $a|$ to a quarter-interval we'd be done. We can do just that with a contraction that sends the top of the image of $a|$ to the top of a quarter-interval. The top of the image is $(a|\top)$. The contraction at $(a|\top)^{\bullet}$ sends it to the top of $[-\frac{1}{2}, 0]$.

5. *Diversion:* Łukasiewicz vs. Girard

On the unit interval the formula for $\multimap$ has a prior history as the Łukasiewicz notion of many-valued logical implication. A traditional interpretation of $\Phi \leqq \Psi$ is "Ψ is at least as likely as Φ." Then $\Phi \multimap \Psi$ becomes the "likelihood of Ψ being at least as likely as Φ." [46]

The unit interval viewed as a $\star$-autonomous category (resulting from its scale-algebra structure) reveals Łukasiewicz inference as a special case of Girard's linear logic when we write $\overset{\wedge}{|}$ as $\otimes$ and its "de Morgan dual" $\overset{\vee}{|}$ as $⅋$ ("par"). (The midpoint operation is an example of a "seq" operation—it lies between $\otimes$ and $⅋$.)

If we interpret the truth-values as frequencies (or probabilities) we can not infer, of course, the frequency of a conjunction from the individual frequencies. But we can infer the *range* of possible frequencies. If Φ and Ψ are the individual frequencies then the maximum possible frequency for their disjunction occurs when they are maximally exclusive: if their frequencies add to 1 or less and if they never occur together then the maximal possible frequency of the disjunction is their sum; if their frequencies add to more than 1 then the maximal possible frequency of the disjunction is, of course, 1. That is, the maximum possible frequency for their disjunction is $\Phi ⅋ \Psi$. The minimal possible frequency for their conjunction likewise occurs when they are maximally exclusive and similar consideration yields $\Phi \otimes \Psi$. This works best if we understand that separate observations are made, one for Φ and one for Ψ (hence $\Phi \otimes \Phi$ is the minimal possible frequency that Φ occurs in both observations, $\Phi ⅋ \Phi$ the maximal possible frequency that Φ occurs in at least one of the two observations).

The "additive connectives," likewise, have such an interpretation. The minimal possible frequency for their disjunction occurs when they are minimally exclusive, that is, when the less probable event occurs only when the more probable event occurs, hence the minimal possible frequency for the disjunction is $\Phi \vee \Psi$. Similar computation yields that the maximal possible frequency of their conjunction is $\Phi \wedge \Psi$. The range of frequencies possible for the conjunction is the interval $[\Phi \otimes \Psi, \Phi \wedge \Psi]$ and the range for the disjunction is $[\Phi \vee \Psi, \Phi ⅋ \Psi]$. The midpoint of Φ and Ψ is also the midpoint of $\Phi \otimes \Psi$ and $\Phi ⅋ \Psi$ (using the law of compensation) and the midpoint of $\Phi \wedge \Psi$ and $\Phi \vee \Psi$ (using the forthcoming linear representation theorem of Section 8 (p22–25). Note that we have in descending order:

$$\begin{array}{c} \mathbf{1} \\ \Phi ⅋ \Psi \\ \Phi \vee \Psi \\ \Phi \mid \Psi \\ \Phi \wedge \Psi \\ \Phi \otimes \Psi \\ \mathbf{0} \end{array}$$

When $\Phi \multimap \Psi = 1$ it is possible (just knowing the frequencies of Φ and Ψ) that whenever Φ occurs Ψ will occur. More generally $\Phi \multimap \Psi$ gives the maximal possible probability that a single pair of observations will fail to falsify the hypothesis "if Φ then Ψ." The adjointness lemma, $\Phi \leqq \Psi \multimap \Lambda \quad \Leftrightarrow \quad \Phi \otimes \Psi \leqq \Lambda$, then says that Φ is possibly less frequent than Ψ appearing to imply Λ iff the frequency of the conjunction of Φ and Ψ is possibly less than the frequency of Λ. [47] We constructed the meet operation as $\Phi \otimes (\Phi \multimap \Psi)$. That is, the maximal possible frequency for the conjunction of two events is equal to the minimal possible

[46] We may interpret $\perp$-zooming using the equation $\dot{\Phi} \multimap \Phi = \overset{\vee}{\Phi}$: given a sentence Φ it says that $\overset{\vee}{\Phi}$ is the likelihood that Φ is at least as likely as not. Using the companion equation $\Phi \multimap \dot{\Phi} = \overset{\wedge}{\Phi}$ we see that in the Łukasiewicz interpretation the coalgebra condition ($\overset{\vee}{\Phi} = \top$ *or* $\overset{\wedge}{\Phi} = \perp$) says that for any statement either it or its negation is at least as likely as not.

[47] This would not, of course, be heard as an acceptable sentence in ordinary language. But few translations from the

frequency of the conjunction of another pair of events, the first of which remains the same and the second is the maximal possible frequency of failing to refute the hypothesis that the first implies the second. (Surely someone previously must have observed all this.)

When the coalgebra condition is interpreted we obtain the **interval rule** for linear logic:

$$\Phi \parr \Psi = 1 \quad \text{or} \quad \Phi \otimes \Psi = 0$$

(Either it is possible for one to succeed or it is possible that both fail.) Alternatively we may replace Φ with $\dot{\Phi}$ so that the coalgebra condition becomes

$$\Phi \leqq \Psi \quad \text{or} \quad \Psi \leqq \Phi$$

delivering a theory of linear linear (or planar?) logic.

Missing above are Girard's modal unary operations, *of-course* and *why-not*, which he denoted with a **!** and a **?**. [48] In Section 19 (p43–44) on "chromatic scales" we introduce the (discontinuous) "support" operations on scales. Using chromatic-scale notation one may argue that $!\Phi = \underline{\Phi}$ and $?\Phi = \overline{\Phi}$.

6. *Diversion:* The Final Interval Coalgebra as a Scale

The final interval coalgebra, $\mathtt{I}$, comes equipped, of course, with the two constants $\top$ and $\bot$, and the two zooming operations $\hat{x}$ and $\overset{\vee}{x}$. As previously noted in Section 1 (p5–7) we may define $\dot{x}$ via the unique coalgebra map $\dot{\mathtt{I}} \to \mathtt{I}$ where $\dot{\mathtt{I}}$ is the coalgebra obtained by swapping the two constants and the two zooming operations. The order on $\mathtt{I}$ is definable via the observation that $\dot{x} < y$ iff there is a sequence of zooming operations $(^\wedge, ^\vee)$ that carries x to $\bot$ and y to $\top$.

There is, indeed, a useful interval coalgebra structure on $\mathtt{I} \times \mathtt{I}$ so that its unique coalgebra map to $\mathtt{I}$ is the midpoint operation,[49] but, alas, this coalgebra structure on $\mathtt{I} \times \mathtt{I}$ requires the midpoint operation for the construction of its two zooming operations: $\langle u, v \rangle$ is sent by $\top$-zooming to $\langle \overset{\vee}{u} \overset{\wedge}{|} \hat{v}\, ,\ \hat{u} \overset{\wedge}{|} \overset{\vee}{v} \rangle$ and by $\bot$-zooming to $\langle \overset{\vee}{u} \overset{\vee}{|} \hat{v}\, ,\ \hat{u} \overset{\vee}{|} \overset{\vee}{v} \rangle$. [50]

We must eventually come to grips with the notion of *co-recursion* but will settle now for a quick and dirty proof that for $u, v \in [0, 1]$ the binary expansion of $u|v$ is forced by the scale axioms. Recall the earlier quick and dirty proof. In this case it says that we should iterate (forever) a procedure equivalent to: `If` $u \overset{\vee}{|} v = \top$ `then emit 1 and replace` $u|v$ `with` $u \overset{\wedge}{|} v$ `else emit 0 and replace` $u|v$ `with` $u \overset{\vee}{|} v$. We need, obviously, to expand.

We will use that $u \overset{\vee}{|} v = \top$ iff $\dot{u} \leqq v$ and we will attack the computation of $u \overset{\wedge}{|} v$ and $u \overset{\vee}{|} v$ by using the scale identity. We need a single procedure for the three cases $u|v$, $u \overset{\wedge}{|} v$, $u \overset{\vee}{|} v$.

---

mathematical notation to ordinary language yield acceptable sentences—else who would need the math? Note that the next step (provided by [44], p17) would be the internalization: $\Phi \multimap (\Psi \multimap \Lambda) = (\Phi \otimes \Psi) \multimap \Lambda$. Anyone want to try a translation?

[48] Hollow men pronounce these as *bang* and *whimper*.

[49] The word "useful" is important here. Given any functor T with a final coalgebra $F \to TF$ then for any retraction $F \xrightarrow{x} A \xrightarrow{y} F = 1_F$ there is a (not very useful) coalgebra structure on A that makes y a coalgebra map, to wit, $A \xrightarrow{y} F \to TF \xrightarrow{Tx} TA$.

[50] Imagine stumbling across this use of the scale identity, the initial discovery of which was in answer to a very different question.

Hence we iterate (forever) a procedure that takes an ordered triple $\langle u, s, v\rangle$ as input where u and v are elements of $[0,1]$ and s is an element of the set of three symbols $\{\overset{\vee}{|}, |, \overset{\wedge}{|}\}$.

```
If  s = | then
        if u̇ ≤ v then emit 1; replace ⟨u,|,v⟩ with ⟨u,^|,v⟩.
        else emit 0; replace ⟨u,|,v⟩ with ⟨u,v|,v⟩.
else if s = ^| then if û = ⊥ then emit 0; replace ⟨u,^|,v⟩ with ⟨ǔ,^|,v̂⟩.
        else if v̂ = ⊥ then emit 0; replace ⟨u,^|,v⟩ with ⟨û,^|,v̌⟩.
        else replace ⟨u,^|,v⟩ with ⟨û,|,v̂⟩.
else if ǔ = ⊤ then emit 1; replace ⟨u,v|,v⟩ with ⟨û,v|,v̌⟩.
        else if v̌ = ⊤ then emit 1; replace ⟨u,v|,v⟩ with ⟨ǔ,v|,v̂⟩.
        else replace ⟨u,v|,v⟩ with ⟨ǔ,|,v̌⟩.
```

For a proof that this is forced by the axioms for midpointing note first that $\dot{u} \leq v$ implies $\odot \leq u|v$, hence $u\overset{\vee}{|}v = \top$ which means that the first digit is 1 and the remaining digits are determined by $u\overset{\wedge}{|}v$. For $u\overset{\wedge}{|}v$ we use the scale identity:

$$u\overset{\wedge}{|}v = (\check{u} \overset{\wedge}{|} \hat{v}) \mid (\hat{u} \overset{\wedge}{|} \check{v})$$

When $\hat{u} = \bot$ this becomes:

$$u\overset{\wedge}{|}v = (\check{u} \overset{\wedge}{|} \hat{v}) \mid (\bot \overset{\wedge}{|} \check{v}) = (\check{u} \overset{\wedge}{|} \hat{v})|\bot$$

hence, by the absorbing law, $(u\overset{\wedge}{|}v)^\wedge = \bot$ which means that the first digit is 0 and the remaining digits are determined by $(u\overset{\wedge}{|}v)^\vee = \left((\check{u} \overset{\wedge}{|} \hat{v})|\bot\right)^\vee$ which by the unital law is $\check{u} \overset{\wedge}{|} \hat{v}$. A similar argument holds for the case $\hat{v} = \bot$. If neither $\hat{u}$ nor $\hat{v}$ are $\bot$ we have $\check{u} = \check{v} = \top$ and the scale identity and unital law yield

$$u\overset{\wedge}{|}v = (\top \overset{\wedge}{|} \hat{v}) \mid (\hat{u} \overset{\wedge}{|} \top) = \hat{v} \mid \hat{u}$$

which returns us to the case $s = |$. The dual argument holds for the case $s = \overset{\vee}{|}$.

As previously commented ([7], p5) there are contexts in which $\mathbf{I} \cong \mathbf{I} \vee \mathbf{I}$ is a pushout:

$$\begin{array}{ccc} \bullet & \xrightarrow{\bot} & \mathbf{I} \\ \top\downarrow & & \downarrow \top|x \\ \mathbf{I} & \underset{\bot|x}{\longrightarrow} & \mathbf{I} \end{array}$$

We can use the scale structure on $\mathbf{I}$ to effect that comment. Working in the category whose objects are scales but whose maps are arbitrary functions between them let $f_\bot, f_\top : \mathbf{I} \to S$ be such that $f_\bot(\top) = f_\top(\bot)$. (We're in a full subcaegory of the category of sets; there are no other conditions.) Then $\binom{f_\top}{f_\bot} : \mathbf{I} \to S$ is the map that sends x to $c \triangleleft \left(f_\bot(\check{x}) \mid f_\top(\hat{x})\right)$ where $c = f_\bot(\top) = f_\top(\bot)$.

7. Congruences, or: ⊤–Faces

One of our first aims is to prove that every scale can be embedded in a product of linear scales. Put another way: we wish to find, on any scale, a lot of quotient structures that are linearly ordered. And for that we must get an understanding of quotient structures.

As for any equational theory, the quotient structures of a particular algebra correspond to the "congruences" on that structure, that is, the equivalence relations that are compatible with the operators that define the structure. For some well-endowed theories the congruences correspond, in turn, to certain subsets. Such is the case for scales.[51]

Given a congruence $\equiv$ define its **kernel**, denoted **ker($\equiv$)**, to be the set of elements congruent to $\top$. Clearly, $\equiv$ can be recovered from $\ker(\equiv)$ (because $x \equiv y$ iff both $x \multimap y$ and $y \multimap x$ are in $\ker(\equiv)$). We need to characterize the subsets that appear as kernels.

Borrowing again from convex-set terminology, we say that a subset is a "face" if it is not just closed under midpointing but has the property that it includes any two elements whenever it includes their midpoint. Saying that an element is an extreme point, therefore, is the same as saying that it forms a one-element face. (See Section 41, p95, for a discussion of these definitions.) We will be interested particularly in those faces that include $\top$. Thus we define a subset, $\mathcal{F}$, to be a **⊤-face**, "top-face," if:

$$\top \in \mathcal{F}$$
$$x|y \in \mathcal{F} \quad \text{iff} \quad x \in \mathcal{F} \text{ and } y \in \mathcal{F}$$

Because inverse homomorphic images of faces are faces and because $\{\top\}$ is a face it is clear that $\ker(\equiv)$ is a ⊤-face for any congruence. We need to show that all ⊤-faces so arise.

Given a ⊤-face $\mathcal{F}$ define $\boldsymbol{x \preceq y}$ **(mod $\boldsymbol{\mathcal{F}}$)** to mean $x \multimap y \in \mathcal{F}$ and define $\boldsymbol{x \equiv y}$ **(mod $\boldsymbol{\mathcal{F}}$)** as the "symmetric part" of $\preceq$, that is, $x \equiv y$ iff $x \preceq y$ and $y \preceq x$. It is routine that $x \equiv \top$ iff $x \in \mathcal{F}$.

Clearly $\equiv$ is reflexive (because $x \multimap x = \top$) and it is symmetric by fiat. Transitivity requires a little more. First note that a ⊤-face is an updeal, that is, $x \in \mathcal{F}$ plus $x \le y$ implies $y \in \mathcal{F}$ (immediate from the law of balance). Second, in the dual of the convexity of ⊤-zooming, $\overset{\vee}{u} \,|\, \overset{\vee}{v} \le u \overset{\curlyvee}{|} v$, replace u with $\dot{w}\,|x$ and v with $\dot{x}\,|y$ to obtain:

$$(w \multimap x) \,\big|\, (x \multimap y) \;\le\; \top\big|(w \multimap y) \quad [52]$$

because $(w \multimap x)\big|(x \multimap y) = (\dot{w} \overset{\curlyvee}{|} x)\big|(\dot{x} \overset{\curlyvee}{|} y) \le (\dot{w}\,|x)\overset{\curlyvee}{|}(\dot{x}\,|y) = (\dot{x}\,|x)\overset{\curlyvee}{|}(\dot{w}\,|y) = \odot\overset{\curlyvee}{|}(\dot{w}\,|y) = \top\big|(\dot{w} \overset{\curlyvee}{|} y) = \top\big|(w \multimap y)$. Hence if $w \preceq x$ and $x \preceq y$ then $(w \multimap x)\big|(x \multimap y) \in \mathcal{F}$ forcing $\top\big|(w \multimap y) \in \mathcal{F}$ and, finally, $(w \multimap y) \in \mathcal{F}$.

Thus $\equiv$ is an equivalence relation. It is a congruence with respect to dotting because $w \multimap x = \dot{x} \multimap \dot{w}$, hence $\dot{w} \preceq \dot{x}$ iff $x \preceq w$. In the verification that $y|$ is covariant we used the equation $(y|w) \multimap (y|x) = \top\big|(w \multimap x)$ which quite suffices to show that $w \preceq x$ iff $y|w \preceq y|x$ and consequently that $\equiv$ is a congruence with respect to midpointing. Finally, to see that $\equiv$

[51] Almost all well-endowed theories in nature contain the theory of groups. Two exceptions (besides scales): the theory of division allegories and (its better-known special case) the theory of Heyting algebras.

[52] If both sides of this semiquation are ⊤-zoomed we obtain

$$(w \multimap x) \overset{\wedge}{|} (x \multimap y) \le (w \multimap y).$$

When this semiquation is viewed as a map in a monoidal closed category:

$$(w \multimap x) \otimes (x \multimap y) \to (w \multimap y)$$

its name is the "composition map."

is a congruence with respect to zooming it suffices to show that $w \preceq x$ implies $\hat{w} \preceq \hat{x}$. We may as well show that it implies $\overset{\vee}{w} \preceq \overset{\vee}{x}$ at the same time. The scale identity, in the form $w \multimap x \;=\; (\hat{w} \multimap \hat{x}) \mid (\overset{\vee}{w} \multimap \overset{\vee}{x})$ does just that.

Given an element s in a scale we will need to see how to construct $(\!(s)\!)$ the **principal ⊤-face** it generates, that is, the smallest ⊤-face containing s.

7.1 LEMMA: *The principal ⊤-face $(\!(s)\!)$ is the set of all x such that $s \leqq (\top|)^n x$ for all large n.*

($(\top|)^n$ is the nth iterate of the contraction at ⊤.) Clearly this set includes ⊤ and is closed under midpointing; for the other direction, suppose it includes $x|y$; then from $s \leqq (\top|)^n(x|y)$ we may infer $s \leqq (\top|)^n(x|y) \leqq (\top|)^n(\top|y) = (\top|)^{n+1}y$ for sufficiently large n and y is clearly in the set. For the other direction note that in any quotient where s becomes ⊤ the equality $s \leqq (\top|)^n x$ clearly forces $(\top|)^n x$ to become ⊤, after which n applications of ⊤-zooming will force x itself to be ⊤. That is, if $s \leqq (\top|)^n x$ for any n, then x must be in any ⊤-face that contains s.

8. The Linear Representation Theorem

We wish to prove:

8.1 THEOREM: *Every scale can be embedded in a product of linear scales.*

An algebra (for any equational theory) is said to be **subdirectly irreducible**, or SDI for short, if whenever it is embedded into a product of algebras one of the coordinate maps is itself an embedding. Every algebra (for any equational theory) is embedded in the product of all of its SDI quotients (we will repeat the proof for this case). But first:

8.2 LEMMA: *If a scale is an* SDI *then it is linearly ordered.*

A homomorphism of scales is an embedding iff its kernel is trivial. A scale is an SDI iff the map into the product of all of its proper quotient scales fails to be an embedding. Hence it is an SDI iff the intersection of all non-trivial ⊤-faces is non-trivial. Let $s < \top$ be an element in that minimal non-trivial ⊤-face. Then for every element $a < \top$, its principal ⊤-face, $(\!(a)\!)$, must contain s. Thus an SDI scale has an element $s < \top$ such that for all $a < \top$ it is the case that $a < (\top|)^n s$ for almost all n. (This may be rephrased: a scale is an SDI iff there is a sequence of the form $\{(\top|)^n s\}_n$ cofinal among elements below ⊤.) If x and y are both below ⊤ then clearly $x \vee y < (\top|)^n s$ almost all n, in particular $x \vee y$ is below top. That is, SDI scales satisfy the disjunction property which, as has already been observed in Lemma 4.11, implies linearity via the equation of linearity, $(x \multimap y) \vee (y \multimap x) \;=\; \top$. [53]

The fact that all scales can be embedded in a product of linear scales is now easily obtainable: for each element $s < \top$ use the axiom of choice to obtain a ⊤-face, $\mathcal{F}_s$, maximal among ⊤-faces that exclude s; it is routine that in the corresponding quotient scale the element in the image of s becomes equal to ⊤ in every proper quotient thereof; hence it is, as just argued, linearly ordered. The intersection of all the ⊤-faces of the form $\mathcal{F}_s$ is clearly trivial. (Note that the structure of this proof of the linear representation theorem is forced: if the result is true then necessarily every SDI is linear and the theorem is equivalent to SDI being linear.)

An immediate corollary:

[53] There's a softer proof that uses an easy lemma about principal ⊤-faces: $(\!(x)\!) \cap (\!(y)\!) = (\!(x \vee y)\!)$; if x, y are elements in a scale S such that $x \vee y = \top$ then the map $S \to S/(\!(x)\!) \times S/(\!(y)\!)$ is monic. If, further, S is SDI then either $(\!(x)\!)$ or $(\!(y)\!)$ must be $(\!(\top)\!)$. [Added 2008–12–31]

8.3 CORROLLARY: *Every equation, indeed every universal Horn sentence, true for all linear scales is true for all scales.*

It should be noted that the axiom of choice is avoidable for purposes of this corollary. Given a Horn sentence,

$$(s_1 = t_1) \,\&\cdots\&\, (s_n = t_n) \quad\Rightarrow\quad (u = v)$$

suppose there were a counterexample in some scale, A. The elements used for the counterexample generate a countable subscale, A'. The term $(u \multimap v) \wedge (v \multimap u)$ evaluates to an element $b < \top$. We can construct a $\top$-face $\mathcal{F}$ in A' maximal among those that exclude b without using choice since A' is countable. The image of the counterexample in the linear scale $A'/\mathcal{F}$ remains a counterexample.

When working with a linearly ordered set it is completely trivial that covariant functions automatically distribute with the lattice operations.[54] Hence for all scales we have:

$$\begin{aligned}
x\big|(y \wedge z) &= (x|y) \wedge (x|z)\\
x\big|(y \vee z) &= (x|y) \vee (x|z)\\
\widehat{x \wedge z} &= \hat{x} \wedge \hat{z}\\
\widehat{x \vee z} &= \hat{x} \vee \hat{z}\\
(x \wedge z)^{\vee} &= \overset{\vee}{x} \wedge \overset{\vee}{z}\\
(x \vee z)^{\vee} &= \overset{\vee}{x} \vee \overset{\vee}{z}\\
x \wedge (y \vee z) &= (x \wedge y) \vee (x \wedge z)\\
x \vee (y \wedge z) &= (x \vee y) \wedge (x \vee z)
\end{aligned}$$ [55]

It is now easy to check that if $\diamond$ is any binary operation on a linear scale satisfying the "dilatation equation," $a \diamond (a|x) = x$, and if, further, for any fixed a it is covariant in x, then $a \diamond x = a \triangleleft x$, in particular dilatations are self-dual.[56]

Using the linear representation theorem we obtain a proof for a lemma that we will need later:

8.4 LEMMA: *The image of the central contraction, $\odot|$, is the sub-interval* $[\bot|\odot,\ \odot|\top]$.

We need to show that if $\odot|\bot \leqq x \leqq \odot|\top$ then we can solve for $x = \odot|y$. We remove the existential to obtain a Horn sentence by setting $y = \odot \triangleleft x$. Thus we need to show that in any linear scale $\odot|\bot \leqq x \leqq \odot|\top$ implies $x = \odot|(\odot \triangleleft x)$. Linearity allows us to reduce to the two cases $x \leqq \odot$ and $\odot \leqq x$. Symmetry allows us to concentrate on the case $\odot \leqq x$, hence we can assume $x = \overset{\vee}{x} \,|\, \hat{x} = \top|\,\hat{x}$. From $x \leqq \odot|\top$ we infer that $\hat{x} \leqq \widehat{\odot|\top} = \odot$ hence that

[54] One may, of course, de-trivialize by—instead—establishing the lemma that any function from a linear lattice is covariant iff it is a lattice homomorphism.

[55] The last two equations—the definitions of distributive lattices—are, of course, equivalent for any lattice. Note that the distributivity of a lattice is equivalent with it's having an embedding into the product of its linearly ordered quotients: given elements a, b with a not bounded by b, distributivity implies that the $\{\bot, \top\}$-valued characteristic function of a filter maximal among those that contain a but not b is a lattice homomorphism. Its target is not only linear, it has just two elements. (This also all works for Boolean algebras: thus we teach truth tables.)

[56] As promised, we now have $\big((\dot{a}|\bot)\big|x\big)^{\vee\wedge} = \big((\dot{a}|\top)\big|x\big)^{\wedge\vee}$. There are two other corollaries of interest (see Section 46 (p114–119) for subscorings). First, *any dilatation is definable using just central dilatation:* $a \triangleleft x = \odot \triangleleft \Big(\odot \triangleleft \big((\dot{a}|\odot)\big|x\big)\Big)$ because the one appearance of x is in a covariant position and $\odot \triangleleft \Big(\odot \triangleleft \big((\dot{a}|\odot)\big|(a|x)\big)\Big) = \odot \triangleleft \Big(\odot \triangleleft \big((\dot{a}|a)\big|(\odot|x)\big)\Big) = \odot \triangleleft (\odot|x) = x$. Second, *central dilatation is definable using (twice) any one dilatation:* $\odot \triangleleft x = a \triangleleft \Big(a \triangleleft \big((\odot|a)\big|\dot{x}\big)\Big)^{\bullet}$ because the one appearance of x is in a covariant position and $a \triangleleft \Big(a \triangleleft \big((\odot|a)\big|(\odot|x)^{\bullet}\big)\Big)^{\bullet} = a \triangleleft \Big(a \triangleleft \big((\odot|a)\big|(\odot|\dot{x})\big)\Big)^{\bullet} = a \triangleleft \Big(a \triangleleft \big(\odot\big|(a|\dot{x})\big)\Big)^{\bullet} = a \triangleleft \Big(a \triangleleft \big((a|\dot{a})\big|(a|\dot{x})\big)\Big)^{\bullet} = a \triangleleft \Big(a \triangleleft \big(a\big|(\dot{a}|\dot{x})\big)\Big)^{\bullet} = a \triangleleft (\dot{a}|\dot{x})^{\bullet} = a \triangleleft (a|x) = x$. Hence any one dilatation can be used to construct all other dilatations.

$\hat{x} = \check{\hat{x}} \mid \hat{\hat{x}} = \check{\hat{x}} \mid \bot$ which combines to give $x = \top\big|(\bot|\,\check{\hat{x}}) = (\top|\bot)\big|(\top|\,\check{\hat{x}}) = \odot\big|(\top|\,\check{\hat{x}})$. (So $y = \top|\,\check{\hat{x}}$.) It is routine now that $\odot|(\odot\triangleleft x) = x$. [57]

The fact that $\top$-zooming distributes with meet has an important application: the lattice of congruences is distributive. Recall, first, that in any lattice a set is called a "filter" if it is hereditary upwards and closed under finite meets.[58] By a **zoom-invariant filter** we mean a filter closed under the zooming operations. Since $\bot$-zooming is inflationary a filter is zoom-invariant iff it is closed under $\top$-zooming. An important lemma:

8.5 LEMMA: *A subset of a scale is a $\top$-face iff it is a zoom-invariant filter.*

Because: suppose that $\mathcal{F}$ is a filter invariant with respect to zooming; from $x \wedge y = (x \wedge y) \mid (x \wedge y) \leq x|y$ we know that $\mathcal{F}$ is closed under midpointing; that it is a face follows immediately from $\widehat{x|y} \leq \widehat{x|\top} = x$.

The other direction is an immediate consequence of the facts that zoom-invariant filters are preserved under inverse images of homomorphisms and that any $\top$-face is the inverse image of a one-element zoom-invariant filter, to wit, $\{\top\}$. (It is not hard to give a direct proof: we have already noted (p21) that the law of balance says that a $\top$-face is an updeal, that is, if x is an element of a $\top$-face $\mathcal{F}$, and if $x \leq y$ then $y \in \mathcal{F}$; the law of compensation easily implies that $\hat{x} \in \mathcal{F}$; and if x and y are both in $\mathcal{F}$ we finish with $\widehat{x|y} \leq \widehat{x|\top} = x$ and similarly $\widehat{x|y} \leq y$ hence $\widehat{x|y} \leq x \wedge y$.)

8.6 THEOREM: *The congruence lattice of any scale is a "spatial locale."*

The pre-ordained name for the space in question is the **spectrum** of S, denoted $\mathbf{Spec}(S)$.

First, the lattice of filters in any distributive lattice is itself a distributive lattice and the argument continues to work when we replace "filter" with "zoom-invariant filter" for the simple reason that the set of zoom-invariant filters is closed under arbitrary meets and joins. The main observation (for both proofs) is that the join of filters $\mathcal{F}$ and $\mathcal{G}$ is the set $\{\, x \wedge y : x \in \mathcal{F},\ y \in \mathcal{G} \,\}$. (It is clearly closed under meet and if $x \wedge y \leq z$ then, using distributivity of the lattice, $z = (x \vee z) \wedge (y \vee z)$ where, of course, $x \vee z \in \mathcal{F}$ and $y \vee z \in \mathcal{G}$.)

To see that $(\mathcal{F} \vee \mathcal{G}) \cap \mathcal{H} \subseteq (\mathcal{F} \cap \mathcal{H}) \vee (\mathcal{G} \cap \mathcal{H})$ (the reverse containment holds in any lattice) we note that an arbitrary element in the left-hand side is of the form $x \wedge y$ where $x \in \mathcal{F}$,

[57] Let $\overline{x}$ denote the central dilatation $\odot\triangleleft x$. We could take $\overline{x}$ as primitive and define $\hat{x}$ as $\overline{\overline{(\bot|\odot)}|x}$. There is something to be said for this choice. $\overline{x}$ (unlike $\hat{x}$ and $\check{x}$) appears as an innate operation on almost any graphic calculator. At first glance it looks like we could reduce by one the number of axioms. We would take the single $\overline{\odot|x} = x$ and use the previous footnote to obtain the two unital laws.

The important reason for not using this definition is that the origin of the notion of scales would be belied. But there is another: even when the scale identity is translated into this language the equations are not complete. They do not fix the primitive operation, $\overline{x}$. For a separating example take any scale and define $\overline{x}$ as the standard central dilatation with one exception: redefine $\overline{\top}$ any way one chooses. Thus a further equation—besides the translation of the scale identity—is needed to fix the primitive $\overline{x}$ operation as defined from $\hat{x}$. (Without such, note that there is no way of proving that the primitive operation $\overline{x}$ is covariant, hence no way of showing that $\top$-faces arise from congruences and no way of obtaining the linear representation theorem.)

One could redo the notion of minor scale using $\overline{x}$ as the primitive. After the constant law add the three equations $\overline{\odot|x} = x$, $\overline{\top|(\top|x)} = \top$, and $\overline{\bot|(\bot|x)} = \bot$. The proofs that the elements named by the graded terms are closed under dotting and midpointing remain unchanged. That they are closed under $\overline{x}$ one need only verify $\overline{\top} = \overline{\top|(\top|\top)} = \top$, $\overline{\bot} = \overline{\bot|(\bot|\bot)} = \bot$, $\overline{\top|(\bot|x)} = \overline{(\top|\bot)|(\top|x)} = \top|x$, and, similarly, $\overline{\bot|(\top|x)} = \bot|x$. All of which implies that the previous argument that the theory has a unique consistent equational completion still holds.

[58] It's worth noting—in the context of scales—an alternative definition: $\mathcal{F}$ is a filter if

$$\top \in \mathcal{F}$$
$$x \wedge y \in \mathcal{F} \ \text{ iff } \ x \in \mathcal{F} \text{ and } y \in \mathcal{F}$$

$y \in \mathcal{G}$ and $x \wedge y \in \mathcal{H}$. But the last condition implies that both x and y are in $\mathcal{H}$. Hence, $x \in \mathcal{F} \cap \mathcal{H}$ and $y \in \mathcal{G} \cap \mathcal{H}$, thus $x \wedge y \in (\mathcal{F} \cap \mathcal{H}) \vee (\mathcal{G} \cap \mathcal{H})$.

Distributivity, recall, is quite enough to establish that a lattice of congruences is a locale, that is, finite meets distribute with arbitrary joins. It is always a spatial locale: the points are the "prime" congruences, that is, those that are not the intersection of two larger congruences. Translated to filters: $\mathcal{F}$ is a point if it has the property that whenever $x \vee y \in \mathcal{F}$ it is the case that either $x \in \mathcal{F}$ or $y \in \mathcal{F}$. Put another way, of course, the points of Spec(S) are the linearly ordered quotients of S. We'll show (in Section 42, p95–97) that it is compact normal (but not always Hausdorff). We can obtain (just as in the ancestral subject for spectra) a representation of an arbitrary scale (instead of an arbitrary Noetherian ring without nilpotents) as the scale of global sections of a sheaf of linear scales (instead of domains).[59]

We pause to obtain a "pushout lemma" for scales:

8.7 LEMMA: *Let $A \to B$ be monic and $A \to C$ a quotient map. Then in the pushout*

$$\begin{array}{ccc} A & \to & B \\ \downarrow & & \downarrow \\ C & \to & D \end{array}$$

the map $C \to D$ is monic (and, as in any category of algebras, $B \to D$ is a quotient map).

Because, if we view A as a subscale of B and take $\mathcal{F} = \ker(A \to C)$ then we obtain a $\top$-face of B, to wit, $\mathcal{F}^{\Uparrow} = \{\, b \in B \,:\, \exists_{a \in A}\ a \leqq b \,\}$. It is easy to check that $\mathcal{F}^{\Uparrow}$ is zoom-invariant and that $A \cap \mathcal{F}^{\Uparrow} = \mathcal{F}$. Define $D = B/\mathcal{F}^{\Uparrow}$. The map $A \to B \to D$ has the same kernel as $A \to C$ and we obtain an embedding $C \to D$. It is easily checked to yield a pushout diagram.[60]

9. Lipschitz Extensions and I-Scales

Given an equational theory $\mathbb{T}$, we may say that an extension $\mathbb{T}'$ is "co-congruent" if congruences for the operations in $\mathbb{T}$ remain congruences for all the operations in $\mathbb{T}'$. If all new operations are constant then the extension is automatically co-congruent (*e.g.,* the theory of monoids is a co-congruent extension of the theory of semigroups). A more interesting example is the next step: a monoid congruence is automatically a congruence with respect to the entire group structure (because $x \equiv y$ implies $x^{-1} = x^{-1}yy^{-1} \equiv x^{-1}xy^{-1} = y^{-1}$). [61]

[59] Recall that the space of points of a spatial locale is always "sober" (most easily defined as a space maximal among T_0 spaces with the given locale of open sets). There is often a minimal space, one with the fewest points (the pre-ordained name for this condition is "spaced-out"). For any distributive congruence lattice this minimal space does exist: its elements are the congruences of the SDI quotients. (Could this connection between universal algebra and Stone theory be new?)

[60] This pushout lemma fails in most equational theories. In the category of groups consider the pushout square—as above—where B is the alternating group of order 12, A is the Klein group (the unique subgroup of B of order 4), $A \to B$ its inclusion map, and $A \to C$ an epimorpism where C is a group of order 2. Then the pushout, D, is of order 3. The map $C \to D$ is clearly not epic. A more dramatic example is to enlarge B to the alternating group of order 60; D then collapses to a 1-element group.

[61] There are several similar examples that involve a unary involutory operation that delivers something like an inverse. A lattice-congruence on a Boolean algebra is automatically a congruence for negation: if $x \equiv y$ then $\neg x = \neg x \wedge (y \vee \neg y) \equiv \neg x \wedge (x \vee \neg y) = \neg x \wedge \neg y = (\neg x \vee y) \wedge \neg y \equiv (\neg x \vee x) \wedge \neg y = \neg y$. A ring-congruence on a von Neumann strongly regular ring is automatically a congruence for the "pseudo-inverse" (to wit, the unary operation that satisfies $x^2x^* = x = x^{**}$): if $x \equiv y$ then (using that $xx^* = x^*x$ is a consequence of the axioms) $x^* = x^{*2}x \equiv x^{*2}y = x^{*2}y^2y^* = x^{*2}y(y^2y^*)y^* = x^{*2}y^3y^{*2} \equiv x^{*2}x^3y^{*2} = x^*(x^*x^2)xy^{*2} = x^*x^2y^{*2} = xy^{*2} \equiv yy^{*2} = y^*$. (See Section 46, p114–119 for subscorings.)

A congruence on a acale with respect to midpointing and the two zoom operations is automatically a congruence for dotting: if $u \equiv v$ then $\dot{u} = \odot\!\lhd(\odot\,|\,\dot{u}) = \odot\!\lhd\big((\dot{v}\,|v)\big|\,\dot{u}\big) \equiv \odot\!\lhd\big((\dot{v}\,|u)\big|\,\dot{u}\big) = \odot\!\lhd\big((\dot{v}\,|\,\dot{u})\ \big|\ (u|\dot{u})\big) = \odot\!\lhd\big((\dot{v}\,|\,\dot{u})\ \big|\ \odot\big) = \odot\!\lhd\big((\dot{v}\,|\,\dot{u})\ \big|\ (\dot{v}\,|v)\big) = \odot\!\lhd\big(\dot{v}\,\big|(\dot{u}\,|v)\big) \equiv \odot\!\lhd\big(\dot{v}\,\big|(\dot{u}\,|u)\big) = \odot\!\lhd(\dot{v}\,|\,\odot) = \dot{v}$. (These three sequences of equalities can be cut in half with the observation that it suffices in showing two terms are congruent that one of them is congruent to a term that's invariant when the terms are interchanged.)

But these three examples are misleading. In each case the new operation is unique—when it exists—given the old structure. As we will see such is not the case for all the co-congruent extensions of the theory of scales. (Indeed the potentially most

An extension of the theory of scales is co-congruent iff the equivalence relation determined by any $\top$-face respects the new operations that appear in the extended theory. We will use $\boldsymbol{x} \multimapboth \boldsymbol{y}$ to denote $(x \multimap y) \wedge (y \multimap x)$. A new unary operator f is co-congruent if every $\top$-face $\mathcal{F}$ that contains the element $x \multimapboth y$ also contains the element $fx \multimapboth fy$. If we take $\mathcal{F}$ to be the principal $\top$-face $((x \multimapboth y))$ then for co-congruence to hold we must have, for some integer n,

$$x \multimapboth y \;\; \leqq \;\; (\top|)^n (fx \multimapboth fy)$$

When interpreted on the standard interval this becomes the assertion that f is Lipschitz continuous (with Lipschitz constant $\leqq 2^n$):

$$|fx - fy| \;\; \leq \;\; 2^n |x - y|$$ [62]

If we move to the free algebra on two generators for the extended theory we see that co-congruence requires the existence of an n that works in all models. The argument for unary operations easily extends to all arities. Hence for extensions of the theory of scales we will use the phrase **Lipschitz extension** instead of "co-congruent extension."

The first application is that Lipschitz extensions of the theory of scales all enjoy the linear representation theorem. (More important will be the consequence of the section to come: every consistent Lipschitz extension has an interpretation on the standard interval I.)

If M is a monoid, we understand an ***M*-scale** to be a scale on which M acts. We treat the elements of M as naming unary operations on the scale. Given $m \in M$ and x in the scale we use the usual convention of denoting the values of the corresponding unary operation as mx. We will not require that the M-actions be endomorphisms of the entire scale structure or anything else in particular. In all the cases to be discussed, however, the actions will preserve midpointing and the center:

$$m(x|y) = mx|my \qquad\qquad m\odot = \odot$$

Every scale S is canonically an $\mathbb{I}$-scale determined by the induction scheme (as described in Section 3, p12–13): $\top x = x$, $\bot x = \dot{x}$, $(\top|q)x = x|(qx)$ and $(\bot|q)x = \dot{x}\,|(qx)$. [63]

We will be particularly interested in two cases: when $M = \mathrm{I}$ and when M is the submonoid of rationals in the standard interval. The **theory of I-scales** is obtained by adding for all $r \in \mathrm{I}$ and $q \in \mathbb{I}$ with $q \geqq r$ the equation: [64]

$$q \;\geqq\; r\top$$

and if $q \leqq r$:

$$q \;\leqq\; r\top$$

powerful theorem provides existence proofs—on the standard interval—and the most valued such proofs are precisely those for which there's no uniqueness to force the construction. See p28.) Nor is uniqueness sufficient. A meet semi-lattice has at most one lattice structure but notice that on the four-element non-linear lattice the equivalence relation that smashes the three elements below the top to a single point is a meet- but not a lattice-congruence. A lattice has at most one Heyting-algebra structure but the only non-trivial variety of Heyting algebras in which lattice-congruences are automatically Heyting congruences is the variety of Boolean algebras: any non-Boolean Heyting algebra contains a three-element subalgebra and the equivalence relation on the three-element Heyting algebra that smashes the bottom two elements to a point is a lattice- but not a Heyting-algebra congruence. This works also as a one-variable example: to wit, it is a congruence for the lattice structure but not a congruence for the negated semi-lattice structure (as described in Section 34, p80–82).

[62] On the unit interval $x \,\dot{\multimapboth}\, y$ (the dotting operation applied to $x \multimapboth y$) is $|x-y|$. The dual semiquation of $x \multimapboth y \leq (\top|)^n(fx \multimapboth fy)$ is $(\bot|)^n(fx \,\dot{\multimapboth}\, fy) \leq x \,\dot{\multimapboth}\, y$.

[63] The set of operations that preserve midpointing and $\odot$ form a closed midpoint algebra and we can see this action of $\mathbb{I}$ on arbitrary scales as a consequence of the fact that $\mathbb{I}$ is the initial closed midpoint algebra.

[64] Bear in mind that in the presence of a lattice operation any equality is equational: $x \leq y$ is equivalent with $x = x \wedge y$.

The theory is not as it stands Lipschitz.[65] So we impose the further condition

$$x \multimap\!\!\circ\, y \;\leq\; rx \multimap\!\!\circ\, ry$$

There is obviously a unique I-scale structure on the standard interval I.

9.1 LEMMA: *Any consistent theory of scales may be conservatively extended to include the theory of* I*-scales.*[66]

The "compactness argument" is just what is needed: given a non-trivial model S of a given theory every finite set of equations in the extended theory may be modeled on S itself (for each relevant $r \in \mathrm{I}$ we can find $q \in \mathbb{I}$ whose action on S will satisfy the finite number of equations that involve r) and such suffices for consistency.

It is a consequence of the results in the next section that every non-trivial model of a Lipschitz theory of scales has a quotient isomorphic to I with its unique I-action.

10. Simple Scales and the Existence of Standard Models

10.1 THEOREM: *A scale is simple iff the sequence* $\bot, \top|\bot, \top|(\top|\bot), \ldots, (\top|)^n\bot, \ldots$ *is cofinal among all elements below* $\top$.

The cofinality clearly implies simplicity: if $\mathcal{F}$ is a non-trivial $\top$-face then necessarily there exists $x \in \mathcal{F}$, $x \neq \top$ and the cofinality says that $(\top|)^n\bot \in \mathcal{F}$ for some n, hence that $\bot \in \mathcal{F}$, which means—of course—that $\mathcal{F}$ is entire. Conversely, a simple algebra is necessarily an SDI, hence necessarily linear; but we have much more. We may take the element s used above in the characterization of SDI s to be the element $\bot$. Then, since $\bot$ is included in every non-trivial $\top$-face (because in a simple scale the entire set is the only non-trivial $\top$-face) we know that the sequence $\{(\top|)^n\bot\}$ is cofinal among all elements below $\top$.

A scale satisfying this condition is called, of course, **Archimedean**.[67]

10.2 LEMMA: *A scale is simple iff between any two elements there is a constant (that is, an element from the initial subscale).*

One direction is immediate: if between $\top$ and any $x < \top$ there is some constant, then there must be one of the form $(\top|)^n\bot$. The other direction requires a little work. Given $u < v$, let n be minimal such that $v \multimap (\top|)^{n+1}\bot$, The argument requires induction on n. If $n = 0$, that is if $v \multimap u < \odot$ then the two semiquations $u = \top \multimap u \mathrel{\dot\leq} v \multimap u$ and $\dot v = v \multimap \bot \mathrel{\dot\leq} v \multimap u$

[65] Order the polynomial ring $\mathbb{R}[\varepsilon]$ by stipulating $P(\varepsilon) \geq 0$ iff $P(1/n) \geq 0$ for all sufficiently large n. Its standard interval is a scale and the map that sends $P(\varepsilon)$ to $P(2\varepsilon)$ is a scale-automorphism thereon, hence easily satisfies all equations so far (with $r = 1$). It is not Lipschitz.

[66] A standard equational-theory consequence is that any scale is embedded in its I-scale reflection: given a scale S use the equational theory obtained by adding to the theory of scales a constant for each element of S and adding as equations all the variable-free equations in these constants that hold for S.

[67] There are those who say that the property should be known as "Eudoxian." Euclid wrote about it in the *Elements* and Proclus said that the idea was due to one Eudoxus but a case may be made for Archimedes. Euclid's Definition 5 of Book V:

> Magnitudes are said to be in the same ratio, the first to the second and the third to the fourth, when, if any equimultiples whatever are taken of the first and third, and any equimultiples whatever of the second and fourth, the former equimultiples alike exceed, are alike equal to, or alike fall short of, the latter equimultiples respectively taken in corresponding order.

But Archimedes in his *Quadrature of the Parabola* does a better job of isolating the salient point:

> The excess by which the greater of (two) unequal areas exceeds the less, can by being added to itself be made to exceed any given finite area.

Which is how in the absence of the word *positive* one states that any positive quantity when repeatedly doubled becomes arbitrarily large. (Surely such is equivalent to the assertion that any quantity when repeatedly halved becomes arbitrarily small.) Archimedes, note, did not actually claim originality in this; immediately after the line quoted above he writes:

> The early geometers have also used this lemma.

(consequences of the contravariance of $\multimap$ in the first variable and covariance in the second) yield $u < \odot < v$ and we are done. If $n > 0$ we consider the three cases: $u < v \leqq \odot$, $u < \odot < v$ and $\odot \leqq u < v$. In the 1[st] case, we have $\hat{u} = \bot = \hat{v}$, hence (using the scale identity) $v \multimap u = (\check{v} \multimap \check{u}) \mid (\hat{v} \multimap \hat{u}) = (\check{v} \multimap \check{u})|\top$ and we obtain $\check{v} \multimap \check{u} < (\top|)^n\bot$. Hence by the inductive assumption there is a constant $r \in \mathbb{I}$ such that $\check{u} < r < \check{v}$, thus for $q = r|\bot$ we have $u = \check{u}\,|\,\hat{u} = \check{u}|\bot < q < \check{v}|\bot = \check{v}\,|\,\hat{v} = v$. In the 2[nd] case we take, of course, $q = \odot$. In the 3[rd] case we use the 1[st] case to obtain $\dot{v} < q < \dot{u}$ and finish with $u < \dot{q} < v$.

There are two remarkable facts. The first is that there are many simple scales, so many that every non-trivial scale has a simple quotient: use Zorn's lemma on the set of $\top$-faces that do not contain $\bot$.

The second is that there are very few simple scales. Because there is a constant between any two elements we know that elements are distinguished by which constants appear below (or for that matter, above) them, hence there can not be more elements in a simple scale than there are sets of constants: therefore a simple scale has at most $2^{\aleph_0}$ elements.[68]

Since there is no flexibility on what homomorphisms do to constants:

10.3 PROPOSITION: *Given a pair of simple scales there is at most one map from the first to the second.*

Thus the full subcategory of simple scales is a pre-ordered set. It has a maximal element and the name of that maximal element is the closed interval, $\mathtt{I}$. Non-constant maps from simple scales are embeddings, hence every simple scale is uniquely isomorphic to a unique subscale of $\mathtt{I}$. (And for an algebraic construction of $\mathtt{I}$ take any simple quotient of a coproduct over the family of all simple scales.)

Combining the two remarkable facts we obtain

10.4 THEOREM: *Every non-trivial scale—indeed, any non-trivial model of any Lipschitz extension of the theory of scales—has a homomorphism to* $\mathtt{I}$.

One quick corollary: add any set of constants to the theory of scales and any consistent set of equations thereon. Necessarily there is an interpretation for all the constants in $\mathtt{I}$ satisfying all the equations. (Recall that if all the new operations in an extension of the theory of scales are constants then it is automatically a Lipschitz extension.) As an example, adjoin just one constant, a, and a maximal consistent set of axioms of the form $q \multimap a = \top$ and $a \multimap q = \top$ where the qs are restricted to constants, (that is, names of elements of $\mathbb{I}$). Such a maximal consistent extension is, of course, called a "Dedekind cut" and this quick corollary of the standard models theorem (to wit, that any such set of conditions can be realized in $\mathtt{I}$) is, of course, called "Dedekind completeness."

Because every consistent Lipschitz theory of scales can be enlarged to a consistent Lipschitz theory of $\mathtt{I}$-scales (Section 9, p25–27) we obtain, as an immediate consequence of the existence of simple quotients of non-trivial $\mathtt{I}$-scales:

10.5 THEOREM: ON THE EXISTENCE OF STANDARD MODELS
Every consistent Lipschitz extension of the theory of scales has an interpretation on $\mathtt{I}$.

An immediate corollary (logicians would call it a "compactness theorem"): *A Lipschitz extension has a model on* $\mathtt{I}$ *iff each finite set of its equations has a model on* $\mathtt{I}$.

[68] Compare with the theory of groups: no quotient of the rational numbers—viewed as a group under addition—is simple. On the other hand there are simple groups of every infinite cardinality; indeed, every group can be embedded in a simple group. (For infinite G, compose its Cayley representation with the quotient map that kills all permutations that move fewer elements than the cardinality of G.)

Note that every non-trivial linear scale has a unique map to $\mathtt{I}$ (the kernel of the map is the $\top$-face of all elements larger than all $(\top|)^n\bot$). We may rephrase this: consider the category of non-trivial scales that satisfy the coalgebra condition, to wit, $\hat{x} = \bot$ or $\overset{\vee}{x} = \top$ for all x. This category has a final object. (This fact is, of course, much much weaker than the usual statement of the coalgebraic characterization of $\mathtt{I}$, but it is comforting to see it arise in such a purely algebraic manner.)

Since each simple scale has a unique map to $\mathtt{I}$:

10.6 LEMMA: *The maps from a scale to* $\mathtt{I}$ *are known by their kernels.*

Later we will use the fact that the **maximal–$\top$-face spectrum** of A, denoted $\mathbf{Max}(A)$ is canonically equivalent with the set of maps $(A, \mathtt{I})$. One immediate application:

10.7 THEOREM: *The standard interval,* $\mathtt{I}$*, is an injective object in the category of scales.*

Because, given a subscale S' of S and an $\mathtt{I}$-valued map f' from S' we seek an extension to all of S. Let $\mathcal{F} \subseteq S$ be the $\top$-face generated by $\ker(f')$, (that is, the result of adding all elements in S larger than an element in $\ker(f')$) and note that it remains a proper $\top$-face, hence $S/\mathcal{F}$ is non-trivial and we may choose a map $S/\mathcal{F} \to \mathtt{I}$. Define f to be $S \to S/\mathcal{F} \to \mathtt{I}$. The kernel of $S' \to S \to \mathtt{I}$ must, of course, contain $\ker(f')$. But $\ker(f')$ is maximal, hence f' has the same kernel as $S' \to S \to \mathtt{I}$ and as we just noted, maps to $\mathtt{I}$ are known by their kernels: thus f is an extension of f'.[69]

Bear in mind that all these special properties of $\mathtt{I}$ are maintained for any Lipschitz extension of the theory of scales.

Following the language of ring theory we define the **Jacobson radical** of a scale to be the intersection of all its maximal $\top$-faces and we say that a scale is **semi-simple** if its Jacobson radical is trivial. (The name used in the theory of convex sets for maximal proper faces is "facet," hence we could say that the Jacobson radical is the intersection of all the "top-facets." Doing so, of course, means that one must not be bothered by etymology.) A scale is semi-simple iff its representations into simple quotients are collectively faithful. (Hence, a better term for both rings and scales would have been "residually simple.")

10.8 THEOREM: *A scale is semi-simple iff* $\sup_n (\top|)^n\bot = \top$.

In any simple scale the cofinality of $(\top|)^n\bot$ implies the weaker condition that its supremum is $\top$ and such remains the case in any cartesian product of simple scales. To see that the condition implies semi-simplicity we need to show that it implies for every $x < \top$ that there is a simple quotient in which x remains less than $\top$. The condition tells us that there is n such that the equality $(\top|)^n\bot < x$ fails, that is. $((\top|)^n\bot) \multimap x < \top$. By the linear representation theorem we may find a linear quotient in which that failure is maintained, that is, $((\top|)^n\bot) \multimap x$ remains below $\top$, hence in which $x \leqq (\top|)^n\bot$. Now take its (unique) simple quotient.

10.9 LEMMA: *The Jacobson radical of a scale is the set,* $\mathcal{R}$*, of all* x *such that* $(\top|)^n\bot \leqq x$ *for all* n.

It is clear $\mathcal{R}$ is in the kernel of every simple quotient. For the converse we need to show that $((\bot|)^m\top) \multimap x \equiv \top \pmod{\mathcal{R}}$ for if all m, then $x \in \mathcal{R}$. But $((\bot|)^m\top) \multimap x \equiv \top \pmod{\mathcal{R}}$ says, of course, that $((\bot|)^m\top) \multimap x \in \mathcal{R}$ hence $(\top|)^n\bot \leqq ((\top|)^m\bot) \multimap x$ all n. In particular $(\top|)^n\bot \leqq ((\top|)^n\bot) \multimap x$ all n. Use the adjointness lemma to obtain $\widehat{(\top|)^n\bot} \leqq x$ all n, hence that $(\top|)^{n-1}\bot \leqq x$ all n, which, of course, says $x \in \mathcal{R}$.

[69] The injectivity of $\mathtt{I}$ is quite enough to yield the existence of a map to $\mathtt{I}$ from every non-trivial scale since we know that any non-trivial scale contains a copy of at least one scale with a map to $\mathtt{I}$. to wit, $\mathbb{I}$.

Our definition of the Jacobson radical as the intersection of all the maximal $\top$-faces relied on the axiom of choice. But note that this construction of the Jacobson radical is choice-free. (So it would have been better—with both rings and scales—to use the choice-free construction as the definition.)

11. A Few Applications

For the most algebraic construction of $\mathtt{I}$, take the co-product of all one-generator simple scales and reduce by the Jacobson radical. Put another way: start with a freely generated scale and for each generator reduce by a maximal consistent set of relations that involve only that generator (and, of course, the primitive constants of the theory of scales). Do this in such a way that every possible such set of relations appears for at least one generator. Now take any simple quotient. It is necessarily a copy of $\mathtt{I}$. But we do not need the axiom of choice: there is only one simple quotient and it is the result of reducing by the Jacobson radical. (We are in a case where semi-simplicity implies simplicity.)

Footnote [2] (p4) suggested a way of handling limits of sequences in $\mathtt{I}$. Let's redo it, this time without using the axiom of choice. Again let $\mathtt{I}^{\mathbb{N}} = \prod_{\mathbb{N}} \mathtt{I}$ denote the scale of all sequences in $\mathtt{I}$. The first step is to identify sequences that agree almost everywhere to obtain the quotient scale $\mathtt{I}^{\mathbb{N}}/\mathcal{E}$ (where $\mathcal{E}$ is the $\top$-face of all sequences that are eventually constantly equal to $\top$). The next step (a step we could not take before) is to reduce by the Jacobson radical. $(\mathtt{I}^{\mathbb{N}}/\mathcal{E})/\mathcal{R}$. As already observed, there is never more than one map from $\mathtt{I}$ to a semi-simple scale hence the map $\mathtt{I} \overset{\Delta}{\to} \mathtt{I}^{\mathbb{N}} \to (\mathtt{I}^{\mathbb{N}}/\mathcal{E}) \to (\mathtt{I}^{\mathbb{N}}/\mathcal{E})/\mathcal{R}$ is the unique map from $\mathtt{I}$ to $(\mathtt{I}^{\mathbb{N}}/\mathcal{E})/\mathcal{R}$. Let

$$\begin{array}{ccc} C & \to & \mathtt{I}^{\mathbb{N}} \\ \downarrow & & \downarrow \\ \mathtt{I} & \to & (\mathtt{I}^{\mathbb{N}}/\mathcal{E})/\mathcal{R} \end{array}$$

be a pullback where we view both horizontal maps as inclusions. C is the subscale of all convergent sequences. The (vertical) map $C \to \mathtt{I}$ is the unique map from C to $\mathtt{I}$ that respects almost-everywhere equivalence. Its standard name is "$\lim\limits_{n \to \infty}$".

We said in Section 0 (p3–4) that $C \subseteq \mathtt{I}^{\mathbb{N}}$ could be defined as the joint equalizer of all the closed midpoint maps from $\mathtt{I}^{\mathbb{N}}/\mathcal{E}$ to $\mathtt{I}$. Since pullbacks of equalizers are equalizers it suffices, obviously, to show that the (unique) map $\mathtt{I} \to (\mathtt{I}^{\mathbb{N}}/\mathcal{E})\mathcal{R}$ is such a joint equalizer. Define the **simple part** of any semi-simple scale to be the joint equalizer of all maps to $\mathtt{I}$; since, by definition, those maps are jointly faithful any map from the simple part to $\mathtt{I}$ is necessarily an embedding and the simple part is, indeed, simple; conversely any simple subscale has a unique map to $\mathtt{I}$, hence is in the simple part. All of which says that any map from $\mathtt{I}$ to a semi-simple scale, *e.g.*, the diagonal map from $\mathtt{I}$ to $\mathtt{I}^{\mathbb{N}}$ followed by the quotient map to $(\mathtt{I}^{\mathbb{N}}/\mathcal{E})\mathcal{R}$, is automatically its simple part.[70]

[70] In Section 0 (p3–4) we made the joint-equalizer assertion not for scale maps but for closed midpoint maps. For a proof, let $s_1, s_2, \ldots,$ be a sequence in $\mathtt{I}$. Note first that if $q \in \mathbb{I}$ is an upper bound for almost all s_n, then for any closed midpoint map, $f : \mathtt{I}^{\mathbb{N}}/\mathcal{E} \to \mathtt{I}$ we have $f(s) \leq q$, hence $f(s) \leq \limsup s$ and similarly $f(s) \geq \liminf s$. In particular, for any convergent s we have $f(s) = \lim s$. Next, if a is an accumulation point of s then we may take infinite $\mathbb{N}' \subseteq \mathbb{N}$ such that s restricted to $\mathbb{N}'$ converges to a. The inclusion map $\mathbb{N}' \to \mathbb{N}$ induces a scale map $\mathtt{I}^{\mathbb{N}}/\mathcal{E} \to \mathtt{I}^{\mathbb{N}'}/\mathcal{E}$ that carries s to a convergent sequence. The axiom of choice gives us an $\mathtt{I}$-valued scale map, hence any accumulation point of s appears as a value of an $\mathtt{I}$-valued scale map—*a fortiori*, a closed midpoint map—from $\mathtt{I}^{\mathbb{N}}/\mathcal{E}$ to $\mathtt{I}$. Since $\mathtt{I}$-valued closed midpoint maps are closed under convex combinations we may conclude that the closed interval $[\liminf s, \limsup s]$ is the set of all such values. (But the only values of the scale maps are the accumulation points: if b is not an accumulation point we may chose $\ell < b < u$ such that $s_n \notin (\ell, u)$ for almost all n; then $(s_n \multimap \ell) \vee (u \multimap s_n) = \top$ for almost all n; hence $(f(s) \multimap \ell) \vee (u \multimap f(s)) = \top$ and we may conclude $f(s) \notin (\ell, u)$. So every closed subset of $\mathtt{I}$ so appears: surely it's an old exercise that any separable closed subset in any space is the set of accumulation points of a sequence in that space.)

Even before we took limits of sequences we defined $\mathcal{C}(G)$ as the set of continuous maps and promised that " In Section 11 (p30–31) we will obtain a totally algebraic definition." The scale of uniformly continuous $\mathtt{I}$-valued maps on a uniform space, X, is easier to construct. Consider, first, $\mathtt{I}^{X\times X}$ the scale of all functions from $X\times X$ to $\mathtt{I}$. Let $\mathcal{E}$ be the $\top$-face of those functions that are equal to $\top$ on some element of the given uniformity and let $\mathcal{R}$ be the Jacobson radical of $\mathtt{I}^{X\times X}/\mathcal{E}$. The pair of projection maps from $X\times X$ to X induces a pair of maps from $\mathtt{I}^X$ to $\mathtt{I}^{X\times X}$ which, in turn, yields—by composition—a pair of maps from $\mathtt{I}^X$ to $(\mathtt{I}^{X\times X}/\mathcal{E})/\mathcal{R}$. *The scale of uniformly continuous* $\mathtt{I}$*-valued maps on* X *is the equalizer of this final pair of maps.*

For ordinary continuity let X be a topological space and $\mathtt{I}^X = \prod_X \mathtt{I}$ denote the set of all functions (continuous or not) from X to $\mathtt{I}$. For $x \in X$ identify functions that agree on a neighborhood of x to obtain $\mathtt{I}^{\mathbb{N}}/\mathcal{E}_x$ (where $\mathcal{E}_x$ is the $\top$-face of all sequences that are equal to $\top$ on some neighborhood of x). Then reduce by the Jacobson radical $(\mathtt{I}^{\mathbb{N}}/\mathcal{E}_x)/\mathcal{R}$. Let

$$\begin{array}{ccc} C_x & \rightarrow & \mathtt{I}^X \\ \downarrow & & \downarrow \\ \mathtt{I} & \rightarrow & (\mathtt{I}^X/\mathcal{E}_x)/\mathcal{R} \end{array}$$

be a pullback where $C_x \rightarrow \mathtt{I}^X$ is an inclusion. C_x is the subscale of all functions from X to $\mathtt{I}$ that are continuous at x, or, put another way:

11.1 PROPOSITION: *The functions in* $\mathtt{I}^X$ *that are continuous at* x *is the pullback of the simple part of* $(\mathtt{I}^X/\mathcal{E}_x)/\mathcal{R}$.

Finally:

$$\mathcal{C}(X) \;=\; \bigcap_{x\in X} C_x \;\subseteq\; \mathtt{I}^X \qquad [71]$$

12. Non–Semi-Simple Scales and the Richter Scale

The **Richter scale**, R, is defined as the interval from $\langle -1, 0\rangle$ through $\langle 1, 0\rangle$ in the lexicographically ordered $\mathbb{D}$-module $\mathbb{D}\oplus\mathbb{D}$. [72] The set-valued "Jacobson-radical functor" is represented by the Richter scale, that is, the elements of the Jacobson radical of a scale A are in natural one-to-one correspondence with the scale maps from R to A. Mac Lane's "universal element" (the most important concept in *Algebra*!) may be taken to be $\langle 1, -1\rangle$: for every element, x, in the Jacobson radical of A there is a unique map that carries $\langle 1, -1\rangle$ to x. [73]

The Richter scale is not, of course, simple. But it just misses. It has just one quotient neither entire nor trivial, to wit, its semi-simple reflection, $\mathbb{I}$. Hence the Richter scale appears as a subscale of every non–semi-simple scale: the necessary and sufficient condition for semi-simplicity is that a scale not contain a copy of the Richter scale.

[71] A quite different way of getting at $\mathcal{C}(X)$ appears in [130] (p64).

[72] If we view $\mathbb{D}\oplus\mathbb{D}$ as the ring $\mathbb{D}[\hbar]/(\hbar^2)$, ordered so that $\hbar$ is "infinitesimal," then the Richter scale is just the standard interval in the "ring of dyadic dual numbers."Its Jacobson radical is the set of all of its pairs of the form $\langle 1, q\rangle$ (note that q is necessarily non-positive).

[73] For a proof, start with the following:

$$\begin{array}{rcl} f\langle 1, -m2^{-n}\rangle & = & (\top|)^n((x\overset{\wedge}{|}\,)^m\top) \\ f\langle 0, -m2^{-n}\rangle & = & \bot|f\langle 1, -m2^{1-n}\rangle \\ f\langle 0, +m2^{-n}\rangle & = & (f\langle 0, -m2^{-n}\rangle)^{\cdot} \\ f\langle q, \pm m2^{-n}\rangle & = & \odot\triangleleft(q|f\langle 0, \pm m2^{-n}\rangle) \end{array}$$

These formulas can be used, at least, for the uniqueness of f, but the fact they describe a scale map requires a bit of work. When we have in hand the representation theorem of Section 20 (p45–47) for the free scale on one generator an easier proof will be available. (And note, in passing, that $\mathrm{Aut}(\mathsf{R}) \cong \mathbb{Z}$.)

A non-simple scale, S, is SDI iff there is an embedding $\mathsf{R} \to S$ such that $\mathsf{R}_\star$ is cofinal in $S_\star$ (where the "lower star" means "remove the top").[74]

It is worth having at our disposal examples of arbitrarily large SDIs. Let J be any ordered set with top and bottom. Consider the $\mathbb{D}$-module of functions from J to $\mathbb{D}$ with "finite support" (that is, the set of functions that are zero almost everywhere). Lexicographically order these functions and define $\top$ as the characteristic map of $-\infty$ (the bottom of J), define $\bot$ as its negation and define the **Scoville scale**, $Sv[\![J]\!]$, to be its interval $[\bot, \top]$. We may regard R as the subscale of $Sv[\![J]\!]$ of all functions whose support is confined to $\pm\infty$ (where $+\infty$ is the top of J). Every element in $Sv[\![J]\!]$ not in this copy of R is less than any element in R's Jacobson radical and any element $x < \top$ in that radical is such that $\{(\top|)^n x\}$ is cofinal among the elements below $\top$ in all of $Sv[\![J]\!]$. The existence of such an element, recall, is equivalent with $Sv[\![J]\!]$ being an SDI (Section 8, p22–25).

13. A Construction of the Reals

Among the many ways of constructing the reals perhaps the nicest is as the set of "germs of midpoint- and $\odot$-preserving self-maps" on $\mathbf{I}$, that is, a real is named by such a map defined on some open subinterval containing $\odot \in \mathbf{I}$; two such partial maps name the same real if their intersection is also such a partial map. For each real there is a canonical name, to wit, the partial map with the largest domain.

The entire ordered-field structure is inherent: 0 is named by the constant map; 1 by the identity map; negation by dotting; $r+s$ is characterized by $(r+s)x = \odot\lhd\big((rx)|(sx)\big)$; multiplication is, of course, defined as composition (with reciprocation obtained by inverting maps); and $r \leqq s$ iff $rx \leqq sx$ for all positive x near $\odot$. The standard interval in this ring has a (unique) scale-isomorphism to $\mathbf{I}$, to wit, the map that sends r to $r\top$ (as defined by the canonical name for r).

To fill in the details we need:

13.1 LEMMA: *Any midpoint-preserving partial self-map on* $\mathbf{I}$ *with an interval as domain is monotonic.*

It suffices to show that if f preserves midpoints then it preserves betweenness. Suppose that $a < b < c$ and that $f(b)$ is *not* between $f(a)$ and $f(c)$. We will regularize the example by replacing f with $\dot{f}$, if necessary, to ensure that $f(a) < f(c)$. Either $f(b) < f(a) < f(c)$ or $f(a) < f(c) < f(b)$. We may—further—replace fx with $\dot{f}\dot{x}$, if necessary, to ensure the latter. It suffices to show that there is another point b' between a and c which is not only a counterexample, as is b, but doubles, at least, the extent to which it is a counterexample, that is, we will obtain the semiquation $f(b') - f(c) \geqq 2(f(b) - f(c))$. This suffices because if we iterate the construction this distance will eventually be greater than the distance from c to $\top$. The construction of b' is by cases:

$$b' = \begin{cases} c\lhd b & \text{if } b > a|c \\ a\lhd b & \text{if } b < a|c \end{cases}$$

In the first case we have that $f(b) = f(c)|f(b')$, hence $f(b') - f(c) = 2(f(b) - f(c))$. In the second case we have $f(b) = f(a)|f(b')$, hence $f(b') - f(c) = (f(b') - f(a)) + (f(a) - f(c)) = 2(f(b) - f(a)) + (f(a) - f(c)) = 2f(b) - (f(a) + f(c)) \geqq 2f(b) - 2f(c)$. [75]

[74] We will see later that all such $\mathsf{R} \to A$ are "essential extensions" as defined in Section 23 (p51–53).

[75] The monotonicity of midpoint-preserving maps requires linear ordering. Consider the non-constant midpoint-preserving map, $\mathbf{I}\times\mathbf{I} \to \mathbf{I}$ that sends $\langle u, v\rangle$ to $u|\dot{v}$. It sends both $\langle\top,\top\rangle$ and $\langle\bot,\bot\rangle$ to $\odot$ and any monotonic map that collapses top

13.2 COROLLARY: *Any midpoint-preserving partial self-map on* $\mathbf{I}$ *with an interval as domain is determined by its values at any two of its points..*

It is easy enough to construct the midpoint operation on reals: given partial midpoint- and $\odot$-preserving maps r and s define $(r|s)x = (rx)|(sx)$. (The medial law is just what is needed to show that $r|s$ preserves midpoints.) The idempotent, commutative and medial laws for real midpointing are automatic. Using that the image of the central contraction, $\odot|$, is the "middle half," to wit, the sub-interval $[\bot|\odot, \odot|\top]$ (Lemma 8.4, p23) we name the real 2 with the map $\odot\triangleleft$ defined on the middle half. (To check that it preserves midpoints it suffices to check that $\odot|(2(x|y)) = \odot|\left((2x)|(2y)\right)$.) We construct $r{+}s$ as $2(r|s)$ and check that it satisfies $\odot|\left((r{+}s)x\right) = (rx)|(sx)$ for all x near $\odot$. We easily verify—using self-distributivity and the cancellation law—the medial law for addition: $\odot(\odot|((r{+}s){+}(t{+}u))) = \odot(|((r{+}s)|(t{+}u)))) = (\odot|(r{+}s))|(\odot|(t{+}u)) = (r|s)|(t|u)$ and similarly $\odot(\odot|((r{+}t){+}(s{+}u))) = (r|t)|(s|u)$. By fiat 0 is a unit for addition. Associativity is then a consequence of the medial law: $(r + s) + u = (r + s) + (0 + u) = (r + 0) + (s + u) = r + (s + u)$. For commutatity of multiplication it suffices to verify it for reals named by contraction at elements of $\mathbf{I}$ (because by repeated central contractions we can reduce any two arbitrary reals to such). For the commutativity of the multiplicative structure on $\mathbf{I}$ it suffices to verify it on an order-dense subset, to wit, $\mathbb{I}$.

Finally, for distributivity: $r(s + t) = r(2(s|t)) = (r2)(s|t) = (2r)(s|t) = 2(r(s|t)) = 2\left((rs)|(rt)\right) = (rs) + (rt)$.

14. The Enveloping $\mathbb{D}$-module [76]

Given a scale, A, we construct its **enveloping $\mathbb{D}$-module**, M, as the direct limit of:

$$A \xrightarrow{\odot|} A \xrightarrow{\odot|} A \xrightarrow{\odot|} \cdots$$

More explicitly, its elements are named by pairs, $\langle x, m\rangle$, where x is an element of A and m is a natural number. The pair $\langle y, n\rangle$ names the same element iff $(\odot|)^n x = (\odot|)^m y$. Addition is defined by $\langle x, m\rangle + \langle y, n\rangle = \langle\left((\odot|)^n x\right) \mid \left((\odot|)^m y\right), m{+}n{+}1\rangle$. It is routine to check that the definition is independent of choice of name. Commutativity is immediate. The zero-element is named by $\langle\odot, 0\rangle$ and it is routine to see that it is a unit for addition. The medial law can be verified by straightforward application of the definitions. Associativity is—once again—a consequence: $(a + b) + c = (a + b) + (0 + c) = (a + 0) + (b + c) = a + (b + c)$. The negation of $\langle x, m\rangle$ is named by $\langle\dot{x}, m\rangle$. Scalar multiplication by $1/2$ sends $\langle x, m\rangle$ to $\langle\odot|x, m\rangle$.

Embed A into M by sending x to $\langle x, 0\rangle$. [77] We order M by $\langle x, m\rangle \leq$

---

and bottom must, of course, be constant. But simplicity—not just linear ordering—is also required: the non-constant map on the Richter scale that sends $\langle x, v\rangle$ to $\langle 0, v\rangle$ preserves midpoints and, again, sends both ends to $\odot$.

[76] As noted, we usually pronounce "$\mathbb{D}$-module" as "dy-module."

[77] We have used so far only that A is a closed symmetric midpoint algebra, which, recall, was defined to be a model of the three equations for midpointing, the one equation for dotting and the (non-equational) Horn condition of cancellation.

With a little more work one may drop dotting and obtain a representation for plain closed midpoint algebras as follows. Given an object A with a binary operation satisfying the idempotent, commutative and medial laws and the Horn condition of cancellation, define a congruence on the cartesian square $A \times A$ by $\langle u, v\rangle \equiv \langle w, x\rangle$ iff $u|x = v|w$. Reflexivity and symmetry are immediate. For transitivity suppose, further, that $\langle w, x\rangle \equiv \langle y, z\rangle$. Then from $u|x = v|w$ and $w|z = x|y$ we may infer $(u|x)\big|(w|z) = (v|w)\big|(x|y)$ hence $(w|x)\big|(u|z) = (w|x)\big|(v|y)$ which by cancellation wields $\langle u, v\rangle \equiv \langle y, z\rangle$. That it is a congruence is immediate: $\langle u, v\rangle \equiv \langle w, x\rangle$ easily implies $\langle y|u, z|v\rangle \equiv \langle y|w, z|x\rangle$. Let $A \times A \to S$ be the quotient structure. The three equations automatically hold in S but the cancellation condition requires verification (if $\langle y|u, z|v\rangle \equiv \langle y|w, z|x\rangle$ then from $(y|u)\big|(z|x) = (z|v)\big|(y|w)$ we may infer $(y|z)\big|(u|x) = (y|z)\big|(v|w)$ and use cancellation in A to yield $\langle u, v\rangle \equiv \langle w, x\rangle$.) Define dotting on S by $\dot{\langle u, v\rangle} = \langle v, u\rangle$ and verify the constant law. Map A into S by choosing an element c and sending $x \in A$ to $\langle x, c\rangle \in S$. The cancellation law must be used one more time to prove that this is a faithful representation.

In [45] (p17) when verifying that $\mathbb{I}$ is the initial scale it was pointed out that there are cyclic closed symmetric midpoint algebras. We may eliminate them by imposing a torsion-freeness condition, to wit, by adding the further Horn conditions

$\langle y, n\rangle$ iff $(\odot|)^n x \leqq (\odot|)^m y$. We obtain a midpoint-isomorphism from A to the set of all elements in M from $\langle\bot,0\rangle$ through $\langle\top,0\rangle$. That is, $\langle\bot,0\rangle \leqq \langle x,n\rangle \leqq \langle\top,0\rangle$ implies there is $y \in A$ such that $\langle x,n\rangle = \langle y,0\rangle$ and that is because the two M-semiquations translate to the two A-semiquations: $(\odot|)^n\bot \leqq x \leqq (\odot|)^n\top$. When we showed that the image of the central contraction is the sub-interval from $\odot|\bot$ through $\odot|\top$, hence that the image of the nth iterate of the central contraction is the sub-interval $[(\odot|)^n\bot, (\odot|)^n\top]$, we showed—precisely—that it is possible to solve for $\langle x,n\rangle = \langle y,0\rangle$.

We may thus infer:

14.1 LEMMA: *Every scale has a faithful representation as a closed interval in a lattice-ordered $\mathbb{D}$-module.* [78]

15. The Semi-Simplicity of Free Scales

15.1 THEOREM: *An equation in the theory of scales ($\mathbf{I}$-scales) holds for all scales iff it holds for the initial scale, $\mathbb{I}$ ($\mathbf{I}$).*

15.2 COROLLARY: *The theory of scales ($\mathbf{I}$-scales) is a complete equational theory.*

We need to show that if an equation in the operators for scales fails in any scale it fails in $\mathbb{I}$. It suffices, note, to find a failure in $\mathbf{I}$ since the operators are continuous—if a pair of continuous functions disagree anywhere on $\mathbf{I}^n$ they must disagree somewhere on $\mathbb{I}^n$. For reasons to become clear later, we will settle here for a failure in between: we will find a failure on the standard rational interval, $\mathbf{I} \cap \mathbb{Q}$. Given an equation in the operators for scales we already know that if there is a counterexample then there is a counterexample in a linear scale and consequently in a closed interval in an ordered $\mathbb{D}$-module. There are only finitely many elements in the counterexample, hence the ordered $\mathbb{D}$-module may be taken to be finitely generated. The ring $\mathbb{D}$ is a principal ideal domain, hence every finitely generated $\mathbb{D}$-module is a product of cyclic modules, to wit, copies of $\mathbb{D}$ or finite cyclic groups of odd order. But the existence of a total ordering rules out the finite cyclic groups. We are thus in an interval in a totally ordered finite-rank free $\mathbb{D}$-module, hence, *a fortiori*, a totally ordered finite-dimensional $\mathbb{Q}$-vector space. We need to move the counterexample into a totally ordered one-dimensional $\mathbb{Q}$-vector space which, of course, will be taken to be $\mathbb{Q}$ with its standard ordering.

We first translate the given counterexample into a set of $\mathbb{Q}$-linear equalities and semiquations. Besides the variables $x_1, x_2, \ldots, x_n$ that appear in the counterexample we treat $\top$ and $\bot$ as variables. The equation that fails is replaced with a strict semiquation, namely, the strict semiquation that results when the given counterexample is instantiated. For each i we add the two semiquations $\bot \leqq x_i \leqq \top$. And, of course, we add the strict semiquation $\bot < \top$. We eliminate the scale operations by iterating the following substitutions (where A and B are terms free of scale operations, that is, are linear combinations of the variables): replace $A|B$ with $(A+B)/2$; replace $\dot{A}$ with $\bot + \top - A$; replace $\hat{A}$ with either $\bot$ or $2A-\top$, whichever is correct for the given counterexample and if $\hat{A} = \bot$ add to the set of conditions to be satisfied the condition $2A \leqq \bot + \top$.

$[(x|)^p y = y] \Rightarrow [x = y]$, one such condition for each odd prime p. Then one may prove that the enveloping $\mathbb{D}$-module will be torsion-free. If one starts with a finitely generated torsion-free symmetric or plain midpoint algebra one ends in a finite-rank free $\mathbb{D}$-module (using that $\mathbb{D}$ is a principal ideal ring). And one may then embed that module into $\mathbb{R}$, all of which yields a completeness result, to wit, every universally quantified first-order assertion about the (symmetric) plain midpoint algebra $\mathbf{I}$ is true for all torsion-free (symmetric) midpoint algebras. Continuity considerations suffice for the same result with $\mathbf{I}$ replaced by $\mathbb{I}$.

[78] One immediate consequence is the generalization to all contractions of the lemma just used about central contractions. That is, the image of the contraction $a|$ on a scale is the subscale from $a|\bot$ through $a|\top$. (Because the semiquations allow one to prove that $a \triangleleft x = 2x - a$.) An equational proof that $a|\bot \leqq x \leqq a|\top$ implies $x = a|(a \triangleleft x)$ must therefore exist but which proof would appear to be quite incomprehensible.

Next, replace each weak semiquation either with an equality or strict semiquation, depending, once again, on which obtains for the given counterexample. We now eliminate the equations by using the standard substitution technique to eliminate for each equation, one variable (and one equation). We thus are given a finite set of strict linear semiquations and we know that there is a simultaneous solution in a totally ordered finite-dimensional $\mathbb{Q}$-vector space. We wish to find a solution in $\mathbb{Q}$.

We know two proofs, one geometric and the other syntactical.[79]

The geometric argument takes place in Euclidean space. We first standardize the strict linear semiquations to a set $\mathcal{A}$ of "positivities," that is a set of linear combinations of the variables to be modeled as positive elements. We are given a totally ordered finite-dimensional $\mathbb{R}$-vector space with an $(n+2)$-tuple of points one for each variable $\bot, \top, x_1, x_2, \ldots, x_n$. We wish to find an orthogonal projection onto a 1-dimensional subspace L so that all of the linear combinations in $\mathcal{A}$ are sent to the same side in L of the origin. To that end, let P be the polytope whose vertices are precisely those linear combinations in $\mathcal{A}$. Our one use of the total ordering on the vector space is the knowledge that P *does not contain the origin.* Given that fact simply take L to be the 1-dimensional subspace through the point in P nearest to the origin. The image of P on L lies on only one side of the origin which, of course, we declare its positive side. The image of the variables on that line give us an $\mathbb{R}$-instantiation as needed.[80]

The syntactical proof uses an induction on the number of variables $\bot, \top, x_1, x_2, \ldots, x_n$. Let $\mathcal{C}$ be the set of strict semiquations that do not involve the variable x_n. Recast each remaining equality in the form either

$$a_\bot \bot + a_\top \top + a_1 x_1 + a_2 x_2 + \cdots + a_{n-1} x_{n-1} \quad < \quad x_n$$

or

$$x_n \quad < \quad a_\bot \bot + a_\top \top + a_1 x_1 + a_2 x_2 + \cdots + a_{n-1} x_{n-1}$$

where the a_is are rational. Let $\mathcal{L}$ be the set of linear combinations that are to be modeled as strictly less than x_n, and $\mathcal{R}$ the set to be modeled as strictly larger that x_n. If $\mathcal{L}$ is empty use the inductive hypothesis to find an instantiation that models all the semiquations in $\mathcal{C}$ and then choose an instantiation of x_n less than all the modeled values of the combinations in the set $\mathcal{R}$. Dually if $\mathcal{R}$ is empty. If neither is empty, model all the semiquations in $\mathcal{C}$ and all semiquations of the form $L < R$ where $L \in \mathcal{L}$ and $R \in \mathcal{R}$. Then model x_n strictly between the largest of the modeled values of the forms in $\mathcal{L}$ and the smallest in $\mathcal{R}$.

A consequence is that the maps from a free scale into $\mathtt{I}$ are collectively faithful because if two terms can not be proved equal then necessarily there is a counterexample in the free scale, hence in $\mathbb{I}$, *a fortiori* in $\mathtt{I}$. That is, there are elements of $\mathtt{I}$ which—when they are used to instantiate the variables—produce different values for the two terms. Such, of course, describes a scale homomorphism from the free scale that separates the two terms. Hence:

15.3 THEOREM: *Every free scale appears as a subscale of a cartesian power of* $\mathtt{I}$.

An immediate corollary is the semi-simplicity of free scales:

15.4 THEOREM: *Every free scale is embedded into the product of its simple quotients.*

Note that the constructions used for counterexamples of equations in $\mathtt{I}$ can as easily be used for counterexamples of universal Horn sentences: the constructions not only preserve strict semiquations but (in the process of elimination of variables) any number of equalities.

[79] I learned the latter from Dana Scott, who claimed he was only following the lead of Alfred Tarski.

[80] For later purposes, we will need to know that there is a $\mathbb{Q}$-instantiation. Since the problem has been reduced to modeling strict semiquations we may use the continuity of the operations to insure rationality, even dyadic rationality.

Thus given a sentence of the form

$$(s_1 = t_1)\ \&\cdots\&\ (s_n = t_n) \ \Rightarrow\ (u = v)$$

with a counterexample anywhere the constructions produce counterexamples in $\mathbf{I}$, indeed, in the rational points in $\mathbf{I}$. (There are Horn sentences true for some non-trivial scales that do not hold for $\mathbf{I}$. A universal Horn sentence is consistent iff it holds for the initial scale, $\mathbb{I}$. An example of such that does not hold for $\mathbf{I}$ is $[x|(x|\bot) = \top|\bot] \Rightarrow [y = z]$.) We may add one non-equational condition, $\overset{\vee}{x} = \top$ *or* $\hat{x} = \bot$ (which for good historical reasons will be called the **coalgebra condition**) that yields a completeness theorem for the entire universally quantified first-order theory: such a sentence is a consequence of the defining equations for scales plus the coalgebra condition iff it true for $\mathbf{I}$.

16. *Diversion:* Harmonic Scales and Differentiation

The theory of **harmonic scales** is given by a binary operation whose values are denoted with catenation, xy, satisfying the equations:

$$\begin{aligned}
\top x &= x = x\top \\
\odot x &= \odot = x\odot \\
x\,(y|z) &= (xy)|(xz) \\
(x|y)\,z &= (xz)|(yz) \\
(u \multimap\!\!\circ\, v)|(x \multimap\!\!\circ\, y) &\leqq \top|(ux \multimap\!\!\circ\, vy)
\end{aligned}$$

We'll refer to the top row as the "unit condition," the next three rows as as the "bilinear condition." The bottom row is, of course, the Lipschitz condition.[81] Standard multiplication is the unique interpretation of these equations on $\mathbf{I}$, hence there is at most one interpretation on any semi-simple scale.

A few lemmas we'll need: 16.1 LEMMA:

$$\begin{aligned}
\dot{u}v &= (uv)^{\bullet} \\
\bot\, v &= \dot{v} \\
\odot|(uv) &= (\odot|u)v
\end{aligned}$$

For the 1st equation use cancellation on $(uv)|(\dot{u}v) = (u|\dot{u})|v = \odot|v = \odot = (uv)|(uv)^{\bullet}$. For the 2nd: $\bot\, v = \dot{\top} v = (\top v)^{\bullet} = \dot{v}$. For the 3rd: $(\odot|u)v = (\odot v)|(uv) = \odot|(uv)$.

The harmonic structure greatly extends our descriptive power. Among many other things it allows us to identify not just any algebraic number in $\mathbf{I}$ but many transcendentals. As just one example, we will see (p38) that if we add a unary operator f satisfying the condition:

$$\left(\,[(fu)^{\bullet}|(f(v)]\ \middle|\ [(fv)(u|\dot{v})]\,\right)^2 \leqq (u|\dot{v})^4$$

then for any $q \in \mathbb{I}$ it is the case that $q(f\odot) \lesssim \odot|(\odot|(f\top))$ is provable (indeed, provable using only the rules for equational logic[82]) iff $q \lesssim \mathrm{e}/4$ (as defined in the standard interval).

[81] The standard interval in any ordered unital ring with $\frac{1}{2}$ satisfies the first 4 equations. Consider again the ring we used before to show the necessity of a Lipschitz condition: the polynomial ring $\mathbb{R}[\varepsilon]$ ordered by stipulating $P(\varepsilon) \geq 0$ iff $P(1/n) \geq 0$ for all sufficiently large n. If we use, instead, the nonstandard multiplication, $(\sum a_n\varepsilon^n) \circ (\sum b_n\varepsilon^n) = \frac{1}{2}\sum(a_0b_n + a_nb_0 - 2^na_nb_n)\varepsilon^n$ then on $\mathbf{I}$ it is Lipschitz but on the standard interval of $\mathbb{R}[\varepsilon]$ it is not.

[82] A proof that one term is equal to another can always be obtained by a sequence of transformations each of which replaces a subterm of the form $f\langle t_1, t_2, \ldots, t_n\rangle$ with one of the form $g\langle t_1, t_2, \ldots, t_n\rangle$ where the ts are arbitrary terms and the equality of $f\langle x_1, x_2, \ldots, x_n\rangle$ and $g\langle x_1, x_2, \ldots, x_n\rangle$ is an axiom.

The harmonic structure allows us to identify values of derivatives. Suppose f is a unary operator on $\mathbb{I}$ and $a, b \in \mathbb{I}$ are constants:

16.2 PROPOSITION: *For any natural number, n:*

$$[(u|\dot{a})^2]^{\boldsymbol{\cdot}} \leqq (\odot|)^n \left[((fu)^{\boldsymbol{\cdot}}|fa) \mid (b(u|\dot{a}))\right] \leqq (u|\dot{a})^2 \ \textit{implies} \ f'a = b.$$

Switching to usual notation, this is saying that if

$$\frac{fu - fa - b(u-a)}{(u-a)^2}$$

is bounded then $f'a = b$. We may rewrite: if there is a constant K such that

$$-K|u-a| \leqq \frac{fu - fa}{u-a} - b \leqq K|u-a|$$

then

$$\lim_{u\to a} -K|u-a| \leqq \lim_{u\to a} \frac{fu - fa}{u-a} - b \leqq \lim_{u\to a} K|u-a|.$$

hence

$$0 \leqq f'a - b \leqq 0$$

Therefore, if f and g are unary operators on $\mathbb{I}$ satisfying the conditions

$$[(u|\dot{v})^2]^{\boldsymbol{\cdot}} \leqq (\odot|)^n \left[((fu)^{\boldsymbol{\cdot}}|fv) \mid ((gv)(u|\dot{v}))\right] \leqq (u|\dot{v})^2$$

we have $f' = g$ and—since g is bounded—that f is Lipschitz. We will need that last fact, not just on simple scales, but for all scales. Even better, g is also Lipschitz:

16.3 LEMMA: *If f and g are unary operations on a scale and $n \in \mathbb{N}$ such that*

$$[(u|\dot{v})^2]^{\boldsymbol{\cdot}} \leqq (\odot|)^n \left[((fu)^{\boldsymbol{\cdot}}|fv) \mid ((gv)(u|\dot{v}))\right] \leqq (u|\dot{v})^2$$

then f and g are Lipschitz.

Because $\odot|$ is a homomorphism with respect to dotting and midpointing we may replace f with $(\odot|)^n f$ and g with $(\odot|)^n g$ and reduce to the case

$$[(u|\dot{v})^2]^{\boldsymbol{\cdot}} \leqq ((fu)^{\boldsymbol{\cdot}}|fv) \mid \big((gv)(u|\dot{v})\big) \leqq (u|\dot{v})^2$$

We may then apply $(\odot|)$ twice more and replace f with $(\odot|)^2 f$ and g with $(\odot|)^2 g$ thus reducing to the case

$$(\odot|)^2[(u|\dot{v})^2]^{\boldsymbol{\cdot}} \leqq ((fu)^{\boldsymbol{\cdot}}|fv) \mid \big((gv)(u|\dot{v})\big) \leqq (\odot|)^2(u|\dot{v})^2$$

As often when working with harmonic scales we resort to the standard notation appropriate to the enveloping $\mathbb{D}$-module (and, in this case, also negate the terms):

$$-\tfrac{1}{4}(u-v)^2 \leqq fu - fv - (gv)(u-v) \leqq \tfrac{1}{4}(u-v)^2$$

Because g was replaced with $(\odot|)^2 g$ we have the additional information $-\frac{1}{4} \leqq g \leqq \frac{1}{4}$, hence

$$-\tfrac{1}{4}|u-v| - \tfrac{1}{4}(u-v)^2 \leqq fu - fv \leqq \tfrac{1}{4}(u-v)^2 + \tfrac{1}{4}|u-v|$$

If we now use the equality $\frac{1}{4}(u-v)^2 \leqq \frac{1}{2}|u-v|$ we obtain

$$-\tfrac{3}{4}|u-v| \leqq fu - fv \leqq \tfrac{3}{4}|u-v|$$

which, of course, says that f is Lipschitz.

The proof for g is easier. After replacing f and g with $(\odot|)^n f$ and $(\odot|)^n g$ and moving to standard notation we have the two semiquations:

$$-(u-v)^2 \leqq fu - fv - (gv)(u-v) \leqq (u-v)^2$$

By switching u and v we also have (obviously):

$$-(v-u)^2 \leqq fv - fu - (gu)(v-u) \leqq (v-u)^2$$

When we add the two rows we obtain:

$$-2(u-v)^2 \leqq (gu - gv)(u-v) \leqq 2(u-v)^2$$

hence

$$|gu - gv| \leqq 2|u-v|$$

Going back to the condition that allows us to identify e/4, it says that in any model on I we have $f = f'$ hence that $fu = Ae^u$ for some constant A and thus $\mathrm{e}f\odot = f\top$ in any model on I. Since f is necessarily Lipschitz we know that in any simple quotient and any $q \in \mathbb{I}$ we have $q(f\odot) \lesseqgtr \odot|(\odot|(f\top))$ iff $q \lesseqgtr \mathrm{e}/4$, hence such is the case in any linear quotient. The linear representation theorem says, therefore, that in any model, linear or not, and any $q \in \mathbb{I}$ we have $q(f\odot) \lesseqgtr \odot|(\odot|(f\top))$ iff $q \lesseqgtr \mathrm{e}/4$. We can infer more: let F be the initial algebra for the theory of harmonic scales with a unary operator satisfying the condition we used to identify e/4 plus an equation to fix the value of $f\odot$; then its semi-simple reflection, $F/\mathcal{R}$, is simple (to wit, the smallest harmonic subscale of I closed under the action of f). It suffices to show that $F/\mathcal{R}$ is its own simple part: but for any element, a, we have either $q \leqq a$ or $a \leqq q$ for all $q \in \mathbb{I}$, forcing a to be in the equalizer of all maps to I, that is, forcing a to be in the simple part.[83]

A function may have a derivative at a without $[-fu + fa + (f'a)(u-a)]/(u-a)^2$ being bounded (*e.g.*, $f(u) = \sqrt[3]{u^4}$ with $a = 0$) but not when the derivative is Lipschitz; we have the converse of the last lemma:

16.4 LEMMA: *If the derivative of* $f : \mathrm{I} \to \mathrm{I}$ *is* $f' : \mathrm{I} \to \mathrm{I}$ *and, if, further,* f' *is Lipschitz then there is* $n \in \mathbb{N}$ *such than*

$$[(u|\dot{v})^2]^{\bullet} \leqq (\odot|)^n \left[((fu)^{\bullet}|fv) \mid ((f'v)(u|\dot{v})) \right] \leqq (u|\dot{v})^2$$

For a proof let

$$L = \frac{1}{2} \inf_{u \neq v} \frac{f'u - f'v}{u-v} \qquad\qquad U = \frac{1}{2} \sup_{u \neq v} \frac{f'u - f'v}{u-v}$$

We wish to prove the two semiquations:

$$L(u-v)^2 \leqq fu - fv - (f'v)(u-v) \leqq U(u-v)^2$$

We'll prove the first semiquation the second can then be obtained by replacing f and f' with their negations. For each $a \in \mathrm{I}$ let

$$h_a(u) = fu - fa - (f'a)(u-a) - L(u-a)^2$$

[83] In standard notation we thus have that there are non-trivial functions f such that $((1 - x + y)fx - fy)^2 \leq (x-y)^4$ for all $x, y \in [-1, 1]$. For any such f it is the case that $f1/f0 = \mathrm{e}$. (Could this really be new?) By using e^{-x^2} instead of e^x we needn't restrict to the standard interval: there are non-trivial real functions f such that $((2x^2 - 2xy + 1)fx - fy)^2 \leq (x-y)^4$ for all $x, y \in \mathbb{R}$. For any such f it is the case that $f0/f1 = \mathrm{e}$.

The first semiquation is equivalent to $0 \leqq h_a$ for arbitrary a. Clearly $h_a a = h'_a a = 0$. It thus suffices to show that a is an absolute minimal point for h_a and for that it suffices to show that $h'_a u \leqq 0$ for $u < a$ and $h'_a u \geqq 0$ for $u > a$. Since $2L \geqq \frac{f'u - f'a}{u-a}$ we have for $u > a$ that

$$h'_a u = f'u - f'a - 2L(u-a) \geqq f'u - f'a - \frac{f'u - f'a}{(u-a)}(u-a) = 0$$

Note that this (very last) semiquation requires $(u - a) \geqq 0$. The argument for $u < a$ can be obtained by using the present result with $f(x)$ and $f'(x)$ replaced with $f(-x)$ and $f'(-x)$.

The last two lemmas yield:

16.5 THEOREM: *$f: \mathrm{I} \to \mathrm{I}$ is differentiable with a Lipschitz derivative $f': \mathrm{I} \to \mathrm{I}$ iff there is $n \in \mathbb{N}$ such that:*

$$[(u|\dot{v})^2]^{\cdot} \leqq (\odot|)^n \left[((fu)^{\cdot}|fv) \mid ((f'v)(u|\dot{v})) \right] \leqq (u|\dot{v})^2$$

Consequently the harmonic structure allows an exploration of differential equations.

In case one wishes to identify π besides e, add to the theory of harmonic scales a unary operator g subject to the condition:

$$\left(2(1|v^2)(gu|(gv)^{\cdot}) \mid (v|\dot{u})\right)^2 \leqq (u|\dot{v})^4$$

Then for any $q \in \mathbb{I}$ it is the case that $q \lesseqgtr (g\top)|(g\bot)^{\cdot}$ is provable iff $q \lesseqgtr \pi/4$, as defined in the standard interval. (The condition implies $g'(v) = (1+v^2)^{-1}$ in any model in the standard interval.)[84] Again, if we add an equation to fix $g\odot$ we may proceed to show that the semi-simple reflection of the free model for these operators is simple.[85]

The increase in expressive power comes, as usual, with a cost: whereas the first-order theory of harmonic I-scales is decidable (as a consequence of Tarski's proof of the completeness of the first-order theory of real closed fields), the addition of further Lipschitz equational structure can make it possible to capture all of first-order number theory.[86]

---

[84] Using standard notation (and twice the g), if $g'v = 2/(1+v^2)$ then using that the derivative of $2/(1+v^2)$ can be seen to lie between -2 and 2 (most easily so on any graphing calculator) we have $|gu - gv - 2(1+v^2)^{-1}(u-v)| \leq (u-v)^2$ for all $u, v \in \mathbb{R}$. If we multiply by both sides by $(1+v^2)$ we obtain $|(1+v^2)(gu-gv) - 2(u-v)| \leq (1+v^2)(u-v)^2 \leq 2(u-v)^2$, which semiquation is quite enough to imply $g'(v) = 2/(1+v^2)$ hence that $g(v) = 2\arctan(v) + g(0)$ and, finally, $g(1) - g(-1) = \pi$.

[85] It may be the case that the free model is semi-simple (hence simple). To show otherwise we need a term provably greater than $(\top|)^n\bot$ all $n \in \mathbb{N}$ but not provably equal to $\top$. The same goes for the free model used for identifying e.

[86] One way is as follows: the set of positive elements of I under ordinary multiplication and a non-standard "addition" characterized by $(x\dot{+}y)(x|y) = \odot|(xy)$ is isomorphic—via reciprocation—to the real half-line $[1, +\infty)$ (with its usual multiplication and addition). We can identify the reciprocals of positive integers by using the differential equation $x^4h''x - 10x^3h'x + (30x^2+1)hx = 0$: the further equations $h(\pm\pi^{-1}) = 0$ and $h'(\pm\pi^{-1}) = 2^{-6}\pi^{-4}$ identify hx as $(x/2)^6 \sin x^{-1}$ [unfortunately the last exponent did not appear in the TAC version]; hence $x \in \mathrm{I}$ is the reciprocal of a positive integer iff $x > 0$ and $h(\pi^{-1}x) = 0$. Given π^{-1} it thus suffices to add three unary operations h, h', h'' and four conditions:

$$h(\pi^{-1}) = 0$$

$$h'(\pi^{-1}) = 2^{-6}\pi^{-4}$$

$$\left(\tfrac{1}{32}x^4h''x\right)^{\cdot} \Big| \left(\tfrac{5}{16}x^3h'x\right) = \left(\tfrac{15}{16}x^2 \Big| \tfrac{1}{32}\right)$$

$$\left(((hu)^{\cdot}|(hv)) \mid ((h'v)(u|\dot{v}))\right)^2 \vee \left(((h'u)^{\cdot}|(h'v)) \mid ((h''v)(u|\dot{v}))\right)^2 \leqq \left(u|\dot{v}\right)^4$$

where (following Section 3, p12–13):

$$\tfrac{1}{32} = \top|(\bot|(\bot|(\bot|(\bot|\odot))))$$

$$\tfrac{5}{16} = \top|(\bot(|\top|(\bot|\odot)))$$

$$\tfrac{15}{16} = \top|(\top|(\top|(\top|\odot)))$$

17. Subintervals

Given scale elements $b < t$ the interval of all elements x such that $b \leq x \leq t$ is, as usual, denoted $[b, t]$. It is, of course, closed under midpointing, but not, in general, under dotting or zooming. It does have an induced scale-structure. Define a new dilatation operation that sends $x, y \in [b, t]$ to $b \vee (x \lhd y) \wedge t$ (in a distributive lattice $b \vee z \wedge t$ is unambiguous when $b \leq t$). For elements in $[b, t]$ this enjoys the characterizing properties of dilatations, to wit, it is covariant in y and undoes contraction at x (the "dilatation equation" in Section 8, p22–25). We then define the zoom operations on $[b, t]$ as the dilatations at b and t. Dotting is obtained by dilitating "into" its center: $\dot{x} = x \lhd (b|t)$.

The verification of the scale axioms is most easily dispatched by using the semi-simplicity of free scales. If there were a counterexample anywhere there would be one in $\mathbb{I}$. It is easy to see that the induced structure on any non-trivial subinterval of $\mathbb{I}$ makes it isomorphic to $\mathbb{I}$. [87]

Consider the example used in [2] (p4). We considered the set, F, of functions from the standard interval to itself, such that $|f(x)| \leq |x|$ for all x. If we view F as a subset of all functions from the standard interval to itself it is an example of a **twisted interval**. It may be described as the set of functions whose values "lie between the identity function and its negation." Given any scale and elements b, t therein we can formalize the notion by defining the twisted interval $[\![\boldsymbol{b}, \boldsymbol{t}]\!]$ as the set of all elements, x, such that in every linear representation it is the case that x is between b and t. This results, easily enough, in the ordinary interval $[b \wedge t, b \vee t]$. But the scale structure we want on $[\![b, t]\!]$ is different: the bottom is to be b not $b \wedge t$ and the top is to be t not $b \vee t$. We simply repeat the construction as for ordinary intervals but with that one change—the new dilatation operator is still obtained by contracting the output of the ambient operation to the ordinary interval $[b \wedge t, t \vee b]$ it being understood that the top and bottom are not the standard endpoints but rather b and t. We know such yields a scale because we know that on every linear quotient it does so (albeit that on some of those linear quotients the order is not the induced order but its opposite).

18. Extreme Points

There is a curious similarity between idempotents in rings and extreme points in scales. First:

18.1 PROPOSITION: *The following are equivalent:*

$$\begin{array}{c} x \text{ is an extreme point} \\ \hat{x} = x \\ \overset{\vee}{x} = x \\ x \vee \dot{x} = \top \\ x \wedge \dot{x} = \bot \\ \exists_v \left[(x \vee v) = \top = (\dot{x} \vee \dot{v})\right] \\ \exists_v \left[(x \wedge v) = \bot = (\dot{x} \wedge \dot{v})\right] \end{array}$$

Because the law of compensation easily shows, first, that fixed-points for either

[87] But the subintervals of $\mathbb{I}$ are not all isomorphic to each other. We constructed isomorphisms between subintervals of the same length at the end of Section 4, p14–17 (actually, we constructed an isomorphism between any interval and the interval of the same length of the form $[b, \top]$). It is easy to see that subintervals are isomorphic if the ratio of their lengths is a power of 2. The odd part of the numerator of the dyadic rational that measures the length is a complete isomorphism invariant: clearly if a pair of subintervals have the same odd part of their length we may construct an isomorphism; for the converse note that universal Horn sentences of the form $[(\overset{\vee}{x|})^n(\bot) = \odot] \Rightarrow [y = z]$ hold for the induced scale-structure of a subinterval of $\mathbb{I}$ iff n does not divide the odd part of its length.

$\top$- or $\bot$-zooming are automatically fixed-points for the other and, second, that extreme points are fixed-points. Conversely, to show that fixed-points are extreme suppose that $x|y$ is a fixed point; the only cleverness needed is $x = x\overset{\vee}{|}\bot \leq x\overset{\vee}{|}y = x|y = x\overset{\wedge}{|}y \leq x\overset{\wedge}{|}\top = x$.

Before dispatching the remaining conditions note that the interval coalgebra condition, $\overset{\vee}{x} = \top$ *or* $\hat{x} = \bot$, known to be equivalent in scales with linearity, implies that there are just the two fixed-points, $\top$ and $\bot$. On a linear scale either equation $x \vee \dot{x} = \top$ or $x \wedge \dot{x} = \bot$ is thus clearly equivalent with x being extreme and therefore a fixed-point, hence such is the case in any scale. And, similarly, on a linear scale the existence of a complement is also clearly equivalent with being a fixed-point and that dispatches the last of the conditions.

We will use $\boldsymbol{\mathcal{B}(S)}$ to denote the set of fixed-points/extreme-points of a scale, S. The set of extreme points of any scale is closed under dotting and the lattice operations and we will regard $\mathcal{B}$ as a (covariant) functor from scales to Boolean algebras.

Recalling that $\mathcal{C}(X)$ denotes the scale of $\mathbb{I}$-valued continuous functions on a Hausdorff space, X, we see that $\mathcal{B}(\mathcal{C}(X)$ is isomorphic to the Boolean algebra of clopens in X. Following ring-theoretic language, we will say that a scale, S, is **connected** if $\mathcal{B}(S)$ has just two elements, $\top$ and $\bot$.

If A and B are scales then in $A \times B$ the elements $\langle \top, \bot \rangle$ and $\langle \bot, \top \rangle$ are a complementary pair of extreme points. Every complimentary pair of extreme points arises in this way: first, note that if e is an extreme point in a scale S then the principal $\top$-face $(\!(e)\!)$ generated by e is the interval $[e, \top]$ (indeed, any subset that is both an interval and a face must have extreme points as endpoints). The quotient structure $S/(\!(e)\!)$ is isomorphic to the induced scale on the interval $[\bot, e]$ via the map that sends x to $e \wedge x$. The map $S \to [\bot, e] \times [\bot, \dot{e}]$ that sends x to $\langle e \wedge x, \dot{e} \wedge x \rangle$ is an isomorphism; its inverse sends $\langle x, y \rangle$ to $x \vee y$. This is, of course, just the analog of Peirce decomposition for central idempotents.[88] (The fact that $e \wedge x$ describes a homomorphism is most easily dispatched using the linear representation theorem. For example, its preservation of midpointing is the Horn sentence $[\,\hat{e} = e\,] \Rightarrow \big[e \wedge (x|y) = (e \wedge x) \mid (e \wedge y)\big]$, a triviality in any linear scale since its only extreme points are $\bot$ and $\top$.)

The atoms of $\mathcal{B}(S)$ thus correspond to the connected components of S. One consequence is the uniqueness of product decompositions. If S is finite product of connected scales $\mathcal{B}(S)$ is finite; its atoms yield the only decomposition into indecomposable products it has. All of this is just as it is for central idempotents in the theory of rings.[89]

If X is totally disconnected then every $\top$-face in $\mathcal{C}(X)$ is generated by the extreme points it contains; the lattice of $\top$-faces is canonically isomorphic to the lattice of filters in $\mathcal{B}(\mathcal{C}(X))$.

In a product of connected scales $\prod_J S_j$ the extreme points are the characteristic functions of subsets of J. An ultrafilter of $\mathcal{B} = 2^J$ generates a maximal $\top$-face of the product. The quotient structure is usually called an **ultraproduct**. It has the wonderful feature that any first-order sentence is modeled by the ultraproduct iff it is modeled by enough S_js, that is, iff the set of j such that the sentence is modeled by S_j is one of the sets in the ultrafilter.[90]

---

[88] I am a terrible speller myself, but the great first American mathematician deserves to have his name spelled correctly. And pronounced correctly—he and his family (including his son C.S.) rhyme it with *terse.*

[89] But there are a few isomorphisms that are not reminiscent of Peirce decomposition. $[\bot, e]$ is isomorphic to $[\dot{e}, \top]$ via the map that sends $x \in [\bot, e]$ to $\dot{e} \vee x \in [\dot{e}, \top]$. The inverse isomorphism sends $y \in [\dot{e}, \top]$ to $e \wedge y \in [\bot, e]$. The product decomposition arising from an extreme point e can be re-described as the isomorphism to $[\bot, e] \times [e, \top]$ that sends x to $\langle e \wedge x, e \vee x \rangle$. Its inverse sends $\langle x, y \rangle \in [\bot, e] \times [e, \top]$ to $x \vee \dot{e} \wedge y$ (necessarily $x \leq y$).

[90] One may show that any linear scale may be embedded in a subinterval (with its induced scale-structure) of an ultrapower of $\mathbb{I}$.

In a twisted interval $[\![b,t]\!]$ let $b \wedge t$ and $b \vee t$ denote the elements as defined in the ambient scale (since b is bottom according to the intrinsic ordering on the twisted interval the use of the intrinsic—instead of the induced—lattice operations would be unproductive). $b \wedge t$ and $b \vee t$ are a complementary pair of extreme points in the twisted interval and the pair yields an isomorphism $[\![b,t]\!] \to [b \wedge t, b]^\circ \times [b, b \vee t]$ where $^\circ$ denotes the **opposite scale**: the one obtained by swapping $\top$- with $\perp$-zooming and top with bottom. But any scale is isomorphic to its opposite via the dotting operation hence $[\![b,t]\!]$ is isomorphic to $[b \wedge t, b] \times [b \wedge t, t]$ via the map that sends $x \in [\![b,t]\!]$ to $\langle((x \triangleleft (t|b)) \wedge b), x \wedge t\rangle$ (using $(x \wedge b) \triangleleft ((b \wedge t)|b) = (x \triangleleft (t|b)) \wedge b$). And that yields the isomorphism from $[\![b,t]\!]$ to $[b \wedge t, b \vee t]$ that sends x to $((x \triangleleft (t|b)) \wedge b) \vee (x \wedge t)$. All of which totally obscures the geometry of the opening section's construction of derivatives.

19. *Diversion:* Chromatic Scales

A **chromatic scale**[91] is a scale with a (non-Lipschitz, indeed, discontinuous) unary **support operation**, whose values are denoted $\boldsymbol{\overline{x}}$, satisfying the equations:

$$\begin{aligned} \overline{\perp} &= \perp \\ \overset{\wedge}{\overline{x}} &= \overline{x} \\ x \wedge \overline{x} &= x \\ \overline{x \wedge y} &= \overline{x} \wedge \overline{y} \end{aligned}$$

Note that the first three equations have a unique interpretation on any connected scale and the 4th equation holds iff the connected scale is linear.[92]

These equations say, in concert:

19.1 LEMMA: *$\overline{x}$ is the smallest extreme point above x.*

(The 3rd equation says, of course, that $\overline{x} \succeq x$; next, if e is an extreme point then $e = e \vee \overline{\perp} = e \vee \overline{e \wedge \dot{e}} = e \vee (\overline{e} \wedge \overline{\dot{e}}) = (e \vee \overline{e}) \wedge (e \vee \overline{\dot{e}}) = \overline{e} \wedge (e \vee \overline{\dot{e}}) \succeq \overline{e} \wedge (e \vee \dot{e}) = \overline{e} \wedge \top = \overline{e}$; [93] third, if e is an extreme point above x then since the 4th equation implies that the support operation is covariant we have $e = \overline{e} \succeq \overline{x}$).

Note that it follows that the support operator distributes not just with meet but with join (its covariance yields $\overline{x} \vee \overline{y} \preceq \overline{x \vee y}$ and the characterization of $\overline{x \vee y}$ as the smallest extreme point above $x \vee y$ yields $\overline{x \vee y} \preceq \overline{x} \vee \overline{y}$). The **co-support**, $\boldsymbol{\underline{x}}$, of x is the largest extreme point below x. It is easily constructable as $\underline{x} = \dot{\overline{\dot{x}}}$.

19.2 THEOREM: *A scale is a simple chromatic scale iff it is linear. Every chromatic scale is semi-simple, that is, any chromatic scale is embedded (as such) in a product of linear scales. The defining equations for the support operator are therefore complete: any equation—indeed any universal Horn sentence—true for the support operator on all linear scales is a consequence of the four defining equations.*

We have already noticed that connected scales have a support operator iff they are linear, so among chromatic scales connectivity and linearity are equivalent. But connected easily

[91] To some extent, chromatic scales are to measurable functions as scales are to continuous. See, in particular, Section 25 (p57).

[92] The 2nd equation becomes redundant if the 3rd equation is replaced with $x \wedge \dot{\overline{x}} = \perp$. See Section 28 (p61–62).

[93] See Section 46, p114–119 for a subscoring.

implies simple: if a congruence is non-trivial then its kernel has an element, x, below $\top$. But then $\underline{x} \neq \top$ hence $\bot = \underline{x} \equiv \underline{\top} = \top$.

If $\mathcal{F} \subseteq \mathcal{B}$ is a filter in the Boolean algebra of extreme points then the $\top$-face it generates in the scale, $\mathcal{F}^{\Uparrow}$, consists of all elements x such that $\underline{x} \in \mathcal{F}$ (because clearly the set of such x is a zoom-invariant filter).

The support operator is discontinuous on the standard interval but among discontinuous operations it appears to play something of a universal role. The Heyting arrow operation, $u \to v$ can be constructed as $\underline{u \multimap v} \vee v$ (the defining equations for the operation hold for this construction on linear scales, hence the representation theorem implies that they hold on all chromatic scales).[94] We observed in Section 5 (p18–19) that Girard's $!\Phi$ is $\underline{\Phi}$ and his $?\Phi$ is $\overline{\Phi}$.

If $\mathcal{F}$ is an ultrafilter, then for extreme points e and e' if $e \vee e' \in \mathcal{F}$ either $e \in \mathcal{F}$ or $e' \in \mathcal{F}$ consequently if $x \vee x' \equiv \top \pmod{\mathcal{F}^{\Uparrow}}$ then either $x \equiv \top$ or $x' \equiv \top$ forcing the quotient scale to be linear. The scale map to the quotient structure clearly preserves the co-support—and hence the support—operation.

Given any $x \neq \top$ we can find an ultrafilter of extreme points excluding $\underline{x}$ hence x remains below $\top$ in the corresponding quotient structure.[95]

Given any scale S and Boolean algebra B let $S[B]$ be the scale generated by S and the elements of B subject to relations that, first, make those elements fixed-points and, second, obey all the lattice relations that obtain in B; then the maps from $S[B]$ to any scale T are in natural correspondence with the pairs of maps, one a scale-homomorphism $S \to T$, the other a Boolean-homomorphism $B \to \mathcal{B}(T)$. For the special case $S = \mathbb{I}$ we have that the functor $\mathbb{I}[-]$ from Boolean algebras to scales is the left adjoint of $\mathcal{B}(-)$ from scales to Boolean algebras.

It is the case that the adjunction map from B to $\mathcal{B}(S[B])$ is an embedding and in the case that S is connected, $B \to \mathcal{B}(S[B])$ is an isomorphism. This and a number of related issues are discussed in Section 38 (p89–91).

[94] The support operator is definable, in turn, from the Heyting operation, indeed, just from the Heyting negation: $\overline{x} = (x \to \bot)$. (Hence the less colorful alternate name, "Heyting scales.") See Section 34 (p80–82).

[95] The analogous material for rings and idempotents is the following equational theory (which I have assumed for at least 40 years must already be known):

Define a support operation on a ring to be a unary operation satisfying:

$$\begin{array}{rcl} \overline{0} & = & 0 \\ (\ \overline{x}^2 & = & \overline{x}\ ) \\ x\overline{x} & = & x \\ \overline{xy} & = & \overline{x}\,\overline{y} \end{array}$$

(The 2nd equation is, in fact, redundant. It is present only to emphasize the analogy with chromatic scales. See [124], p62.)

For one source of examples take any strongly regular von Neumann ring and take $\overline{x} = xx^*$. For a better source note that the first three equations say, in concert, that a connected ring has a unique support operator and it satisfies the 4th equation iff the ring is a domain (that is, a ring, commutative or not, without zero divisors). These equations are complete for such examples: any equation—indeed any universal Horn sentence—true for all domains is a consequence of these equations because any algebra is embedded (as such) into a product of domains. To prove it, first note that if $x^2 = 0$ then $x = x\overline{x} = x\overline{x}\,\overline{x} = x\overline{x^2} = x\overline{0} = x0 = 0$. Any idempotent is central ($(1-e)xe$ is 0 since its square is 0 hence $xe = exe$ and similarly $ex = exe$). For any idempotent e we have $e = \overline{e}$ (because $e = e + (1-e)\overline{0} = e + (1-e)\overline{(1-e)e} = e + (1-e)\overline{(1-e)}\,\overline{e} = e + (1-e)\overline{e} = e + \overline{e} - e\overline{e} = e + \overline{e} - e = \overline{e}$). (See Section 46 (p114–119) for a subscoring.) The equations characterize $\overline{x}$ as the smallest element in the Boolean algebra of idempotents that acts like the identity element when multiplied by x. If $\mathfrak{A}$ is an ideal in that Boolean algebra then the ideal it generates in the ring consists of all elements x such that $\overline{x} \in \mathfrak{A}$ (if $e \in \mathfrak{A}$ then $\overline{ex} = e\overline{x} \in \mathfrak{A}$ and if, further, $e' \in \mathfrak{A}$ then $(ex + e'x') = (e \vee e')(ex + e'x')$ where $e \vee e' = e + e' - ee' \in \mathfrak{A}$ which gives $e \vee e' \geq \overline{ex + e'x} \in \mathfrak{A}$). When $\mathfrak{A}$ is a maximal Boolean ideal then the ring ideal it generates is prime (since xy is in it iff $\overline{x}\,\overline{y}$ is and a maximal Boolean ideal is a prime ideal in the Boolean algebra) hence the corresponding quotient is a domain. The ring-homomorphism down to the quotient is easily checked to be a homomorphism with respect to the support operation. Finally, given any $x \neq 0$ we can find a maximal Boolean ideal containing $1-\overline{x}$ hence x remains non-zero in the corresponding quotient.

20. The Representation Theorem for Free Scales

20.1 THEOREM: *The free scale* ($\mathbf{I}$*-scale*) *on n generators is isomorphic to the scale of all continuous piecewise* $\mathbb{D}$*-affine* (*affine*) *functions from the standard n-cube,* $\mathbf{I}^n$*, to* $\mathbf{I}$.

We need some definitions.

Given a scale, S, let $\boldsymbol{S[x_1,\ldots,x_n]}$ denote the scale that results from freely adjoining n new elements, traditionally called "variables," $x_1,\ldots,x_n$, to S. (The elements of $S[x_1,\ldots,x_n]$ are named by scale terms built from the elements of S and the symbols $x_1,\ldots,x_n$ with two terms naming the same element iff the equational laws of scales say they must.)

The free scale on n generators is thus $\mathbb{I}[x_1,x_2,\ldots,x_n]$. Each of its simple quotients is the image of a unique map to $\mathbf{I}$, hence is determined by where it sends the generators $x_1,x_2,\ldots,x_n$. And, of course, each x_i may be sent to an arbitrary point in $\mathbf{I}$. Thus $\mathrm{Max}(\mathbb{I}[x_1,x_2,\ldots,x_n])$ may be identified with $\mathbf{I}^n$. Because we now have the semi-simplicity of $\mathbb{I}[x_1,x_2,\ldots,x_n]$ we have a faithful representation $\mathbb{I}[x_1,x_2,\ldots,x_n] \to \mathcal{C}(\mathbf{I}^n)$. In this section we will henceforth treat $\mathbb{I}[x_1,x_2\ldots,x_n]$ as a subscale of $\mathcal{C}(\mathbf{I}^n)$

We will show that any function in $\mathbb{I}[x_1]$ is what is traditionally called "piecewise linear," that is, a function $f:\mathbf{I}\to\mathbf{I}$ for which there is a sequence $\bot = c_0 < c_1 < \cdots < c_k = \top$ such that f is an affine function whenever it is restricted to the closed subinterval from c_i through c_{i+1}. We need to generalize the notion to higher dimensions.

We are confronted with a terminological problem. Tradition has it that piecewise affine functions be called "piecewise linear" but we will need both notions. (In Section 26, p58–59, free lattice-ordered abelian groups will be represented as the functions on $\mathbb{R}^n$ that—informally stated—allow a dissection of $\mathbb{R}^n$ into a finite number of polytopal collections of rays on each of which the function is linear.) Free scales lead not to piecewise linear but piecewise affine functions. Informally: a piecewise affine function is one whose domain may be covered with a finite family of closed polytopes on each of which the function is affine.

So let us start at the beginning. An **affine function** from a convex subset of $\mathbb{R}^n$ to $\mathbb{R}$ can be defined as a continuous function that preserves midpoints.[96] (When the domain is all of $\mathbb{R}^n$ such is equivalent to the preservation of **affine combinations**, that is, combinations of the form $ax+by$ where $a+b=1$. Continuity is then automatic.) It is routine, of course, that a function $f:\mathbb{R}^n\to\mathbb{R}$ is affine iff there are constants $a_0,a_1,\ldots,a_n$ such that $f\langle x_1,x_2,\ldots,x_n\rangle = a_0+a_1x_1+\cdots+a_nx_n$. If f preserves the origin, equivalently if $a_0=0$, it is said to be **linear**.

We will say that an affine function is **$\mathbb{D}$-affine**, pronounced "**dy-affine**," if it carries dyadic rationals to dyadic rationals. It is routine that such is equivalent to the a_is all being in $\mathbb{D}$.

For our purposes the simplest—and technically most useful—definition is that $f:\mathbf{I}^n\to\mathbf{I}$ is **continuous piecewise affine** if it is continuous and if there exists an **affine certification**, to wit, a finite family, $\boldsymbol{\mathcal{A}_f} = \{A_1,A_2,\ldots,A_j\}$, of global affine functions, such that $f(x)\in\{A_1(x),A_2(x),\ldots,A_j(x)\}$ all $x\in\mathbf{I}^n$. Given f and an affine certification, $\mathcal{A}_f$, we construct its **canonical polytopal dissection**, $\mathcal{P}$, as follows: starting with the interior of $\mathbf{I}^n$ remove all hyperplanes that arise as equalizers of two members of $\mathcal{A}_f$; the dense open set that remains then falls apart as the disjoint union of a finite family of open convex polytopes; we take $\boldsymbol{\mathcal{P}_f}$ to be the family of closures of these open polytopes.

Given any $P\in\mathcal{P}$ we know that the functions in $\mathcal{A}_f$ nowhere agree on $\mathring{P}$, the interior of P.

[96] One need not require continuity if the target is bounded: we pointed out in Section 13 (p32–33) that if a midpoint-preserving function lies in a closed interval then it is monotonic, in particular, continuous. Hence if f is a midpoint-preserving map from a convex subset, C, of $\mathbb{R}^n$ to $\mathbf{I}$, then f restricted to any slice through C is continuous and perforce preserves affine combinations. But the preservation of affine combinations is, by definition, something that takes place on slices.

We will use the following observation several times: *If a function g is continuous piecewise affine on a connected set and the functions in its certification nowhere agree on that set then g is not just piecewise affine but affine* This observation does not use anything about affine functions other than their continuity: the equalizers of g and the functions in its certification form a finite family of closed subsets that partition the domain, hence each is a component. Thus f agrees with one element of $\mathcal{A}_f$ on $\mathring{P}$ and—by continuity—on all of P. We denote that affine function as $\boldsymbol{A_P}$.

Note that there is unique minimal affine certification for f (given any certification retain only those elements of the form A_P). The resulting canonical polytopal dissection will, in general, be simpler for smaller certifications.[97]

By a **CPDA function** we mean a continuous piecewise affine function whose certification consists only of $\mathbb{D}$-affine functions.

The fact that CPDA functions are closed under the scale operations is easily established: suppose that g is another CPDA function and that $\mathcal{A}_g$ is its affine certification; then the finite family $\{\, A \,|\, A' : A \in \mathcal{A}_f, A' \in \mathcal{A}_g \,\}$ certifies $f|g$; easier is that $\{\, -A : A \in \mathcal{A}_f \,\}$ certifies $\dot{f}$ and $\{-1\} \cup \{\, 2A{-}1 : A \in \mathcal{A}_f \,\}$ certifies $\hat{f}$. Hence $\mathbb{I}[x_1, x_2, \ldots, x_n]$ viewed as a subset of $\mathcal{C}(\mathbf{I}^n)$, consists only of CPDA functions. (Clearly the generators and constants name $\mathbb{D}$-affine functions.) What we must work for is the converse: that every CPDA function from $\mathbf{I}^n$ to $\mathbf{I}$ so appears. It is fairly routine (but a bit tedious) to verify that $-1 \vee A \wedge +1$ [98] is in $\mathbb{I}[x_1, x_2, \ldots, x_n]$ for every $\mathbb{D}$-affine A. [99]

Let $f : \mathbf{I}^n \to \mathbf{I}$ be an arbitrary CPDA function and $\mathcal{P}$ the canonical polytopal dissection for its minimal certification $\mathcal{A}_f$. We will construct for each pair $P, Q \in \mathcal{P}$ a function $f_{P,Q} \in \mathbb{I}[x_1, x_2, \ldots, x_n]$ such that:

$$f_{P,Q}(x) \;\geqq\; f(x) \quad \text{for} \quad x \in P;$$

$$f_{P,Q}(y) \;\leqq\; f(y) \quad \text{for} \quad y \in Q.$$

(Note that it follows that $f_{P,P}(x) = f(x)$ for $x \in P$.)

Then necessarily:

$$f \;=\; \bigvee_P \left(\bigwedge_Q f_{P,Q} \right)$$

[97] Note, though, that canonical polytopal dissections are not necessarily minimal. With $n = 2$ consider the piecewise $\mathbb{D}$-affine function $(x_1 \wedge x_2) \vee 0$. Its minimal certification is, obviously, $\{x_1, x_2, 0\}$ which yields a canonical polygonal dissection with 6 polygons. But there are many (indeed, infinitely many) dissections with only 4 convex polygons.

[98] In any distributive lattice $\ell \vee (a \wedge u) = (\ell \vee a) \wedge u$ whenever $\ell \leqq u$.

[99] For a proof, say that A is "small" if its values on $\mathbf{I}^n$ lie in $\mathbf{I}$, that is, if $-1 \vee A \wedge +1 = A$. The set of small affine functions is clearly closed under dotting and midpointing. We first show that every small $\mathbb{D}$-affine function is in $\mathbb{I}[x_1, x_2, \ldots, x_n]$ and we will do that by induction. Given a small $\mathbb{D}$-affine $f\langle x_1, x_2, \ldots, x_n\rangle = a_0 + a_1x_1 + a_2x_2 + \cdots + a_nx_n$ we will say that it is of type m if all the a_is can be expressed as dyadic rationals with denominator at most 2^m There are only $2n+3$ small $\mathbb{D}$-affine functions of type 0, to wit, $0, \pm 1, \pm x_1, \pm x_2, \ldots, \pm x_n$. Suppose that we have obtained all small $\mathbb{D}$-affine functions of type m. Given f of type $m{+}1$ let $\{\sigma_i\}_i$ and $\{b_i\}_i$ be such that $a_i = \sigma_i b_i 2^{-(m+1)}$ where $\sigma_i = \pm 1$ and b_i is a natural number. The smallness condition is equivalent to the semiquation $b_0 + b_1 + \cdots + b_n \leq 2^{m+1}$ Let $\{c_i\}_i$ and $\{d_i\}_i$ be sequences of natural numbers such that $b_i = c_i + d_i$ each i and $c_0 + c_1 + \cdots + c_n \leq 2^m$, $d_0 + d_1 + \cdots + d_n \leq 2^m$ Define the small $\mathbb{D}$-affine functions $g\langle x_1, x_2, \ldots, x_n\rangle = \sigma_0 c_0 2^{-m} + \sigma_1 c_1 2^{-m} x_1 + \cdots + \sigma_n c_n 2^{-m} x_n$ and $h\langle x_1, x_2, \ldots, x_n\rangle = \sigma_0 d_0 2^{-m} + \sigma_1 d_1 2^{-m} x_1 + \cdots + \sigma_n d_n 2^{-m} x_n$. Then $f = g|h$. (There are many ways of finding the c_is and d_is, none of which seems to be canonical. Perhaps the easiest to specify is obtained by first defining $e_i = 2^m \wedge \sum_{j=0}^{i} b_j$, then $c_i = e_i - e_{i-1}$ and $d_i = b_i - c_i$.)

Finally, given arbitrary $\mathbb{D}$-affine A let m be such that $2^{-m}A$ is small. Then $-1 \vee A \wedge +1 = (\odot\!\!\lhd)^m(2^{-m}A)$.

We take $f_{P,P}$ (of course) to be $-1 \vee A_P \wedge +1$.

To construct $f_{P,Q}$ for $P \neq Q$ let $\mathcal{D}_P$ denote the set of the functions of the form $A_R - A_S$ that are non-negative on P and let $D_1, D_2, \ldots, D_k$ be the functions in $\mathcal{D}_P$ that are non-positive on Q. Define

$$f_{P,Q} \;=\; -1 \vee A_{P,Q} \wedge +1$$

where

$$A_{P,Q} \;=\; A_P + m(D_1 + D_2 + \cdots + D_k)$$

for suitably large integer m.

The two semiquations for the $f_{P,Q}$s are thus reduced to finding $m \geqq 0$ such that:

$$A_{P,Q}(x) \;\geqq\; A_P(x) \quad \text{for} \quad x \in P$$

$$A_{P,Q}(y) \;\leqq\; A_Q(y) \quad \text{for} \quad y \in Q$$

The 1st semiquation holds regardless of $m \geqq 0$. For the 2nd suppose, first, that $y \in Q$ and $(D_1 + D_2 + \cdots + D_k)(y) = 0$; then necessarily $D_1(y) = \cdots = D_k(y) = 0$ and since P is the intersection of closed half-spaces

$$P \;=\; \bigcap_{D \in \mathcal{D}_P} \{\, x : D(x) \geqq 0 \,\}$$

we may conclude that $y \in P$ hence $A_{P,Q}(y) = A_Q(y) = A_P(y)$. If instead $(D_1 + D_2 + \cdots + D_k)(y) < 0$ then for sufficiently large m we obtain $A_{P,Q}(y) = A_P(y) + m(D_1 + D_2 + \cdots + D_k)(y) \leqq A_Q(y)$. Because the functions are affine, we need this semiquation only for the finite set of extreme points, hence we can choose an m that works for all of them. And that completes the construction.

(One immediate application of all this is a construction for the Richter scale that emphasizes its role as the representor for the Jacobson-radical functor. Start with the free scale on one generator, x, and reduce by the $\top$-face $\mathcal{F}$ generated by all elements of the form $((\top|)^n\bot) \multimap x$ (it produces the minimal congruence that forces x into the Jacobson radical). It is easy to verify that $\mathcal{F}$ is the set of all CPDAs that are constantly equal to $\top$ on some non-trivial interval ending at $\top$; the congruence induced by $\mathcal{F}$ thus identifies two CPDAs precisely when they represent the same germ at $\top$. The congruence class of a CPDA f is determined by the value $f(\top)$ and the left-hand derivative of f at $\top$. Note the curious reversal of sign that is needed to establish an isomorphism with our previous construction of the Richter scale: the element $\langle 1, -1\rangle$ corresponds to a CPDA with a *positive* left-hand derivative at $\top$.)

21. Finitely Presented Scales, or: How Brouwer Made Topology Algebraic

A **finitely generated scale, f.p.scale** is, of course, a scale that appears as a quotient of a finitely presented free scale where the kernel is a finitely generated ⊤-face.

By a **closed piecewise $\mathbb{D}$-affine subset** of Euclidean space we mean a subset of the form $h^{-1}(1)$ where h is a continuous piecewise $\mathbb{D}$-affine function. If X is closed piecewise $\mathbb{D}$-affine we will denote the scale of CPDA $\mathtt{I}$-valued functions on X as $\boldsymbol{CPDA(X)}$. (In the last section we showed that $CPDA(\mathtt{I}^n)$ is isomorphic to $\mathbb{I}[x_1, x_2, \ldots, x_n]$.)

21.1 THEOREM: *Given a closed piecewise $\mathbb{D}$-affine subset X of $\mathtt{I}^n$ we obtain a scale homomorphism $CPDA(\mathtt{I}^n) \to CPDA(X)$ that is onto. Moreover $CPDA(X)$ is an f.p.scale, and all f.p.scales so arise.*

As should be expected, we may remove all the "$\mathbb{D}$-"s:

By a **closed piecewise affine subset** of Euclidean space we mean a subset of the form $h^{-1}(1)$ where h is a continuous piecewise affine function. If X is closed piecewise affine we will denote the scale of CPA $\mathtt{I}$-valued functions on X as $\boldsymbol{CPA(X)}$. (In the last section we showed that $CPA(\mathtt{I}^n)$ is isomorphic to $\mathtt{I}[x_1, x_2, \ldots, x_n]$.)

Because the evaluation maps for points in X are thus collectively faithful we will obtain the immediate corollary:

21.2 THEOREM: *All finitely presented scales* ($\mathtt{I}$*-scales*) *are semi-simple.*[100]

First note that a finitely presented scale needs just one relation, equivalently any finitely generated ⊤-face is principal: it suffices to note that the ⊤-face generated by elements $a_1, a_2, \ldots, a_k$ is generated by the single element $a_1 \wedge a_2 \wedge \ldots \wedge a_k$ (one may use $|$ instead of $\wedge$). When we view the free scale on n generators as the scale of CPDA functions on the standard n-cube it is easily seen that a $\mathbb{D}$-affine function, f, is in the ⊤-face generated by h iff $h^{-1}(1) \subseteq f^{-1}(1)$, hence functions f and g are congruent mod $((h))$ iff they behave the same on the closed $\mathbb{D}$-affine set $S = h^{-1}(1)$. If we had the lemma that CPDA functions on closed $\mathbb{D}$-affine subsets of the cube extend to CPDA functions on the entire cube we would be done. (We will, in passing, prove such to be the case since we will show that any such function is given by an element in the f.p.scale, hence is describable by a term in the free scale and any scale and any such term describes a CPDA function on the standard cube, indeed, on the entire Euclidean space.)

So we must redo the previous proof, this time not for the n-cube but for an arbitrary closed $\mathbb{D}$-affine subset thereof. The only serious complication is that the polytopes of interest are no longer all of dimension n. This complication turns out to be mostly in the eye of the beholder.

Let h be a CPDA function with certification $\mathcal{A}_h = \{A_1, A_2, \ldots, A_k\}$, let $X = h^{-1}(1)$ and let f be an arbitrary CPDA function on X with certification $\mathcal{A}_f = \{A'_1, A'_2, \ldots, A'_l\}$. We seek a term in $\mathbb{I}[x_1, x_2, \ldots, x_n]$ that describes f on X. Given $s \in X$ let $\mathcal{D}_s$ be the set of functions of the form $A_i - A_j$ and $A'_i - A'_j$ that are non-negative on s. Note that the negation of some functions in $\mathcal{D}_s$ can also be in $\mathcal{D}_s$ (to wit, all those that are zero on s).

Define

$$P_s = \bigcap_{D \in \mathcal{D}_s} \{ \, x : D(x) \geq 0 \, \}$$

[100] The injectivity of $\mathtt{I}$ implies that this corollary can be strengthened to all "locally f.p.scales," indeed, it suffices for semi-simplicity that each element be contained in an f.p.subscale.

and let $\overset{\circ}{P_s}$ now be—not the interior but—the *inside* of P_s, that is, the points not contained in any of its proper sub-faces.

Note that if two functions in $\mathcal{D}$ agree anywhere on $\overset{\circ}{P_s}$, they agree everywhere on P_s (and if such happens, we know that P_s is of lower dimension than the ambient Euclidean space). For convenience let $\mathcal{A}_s \subseteq \mathcal{A}_h$ be a minimal certification of h on P_s. Using the lemma from the previous section, we know that h is affine on P_s and since $h(s) = 1$ we know that h is constant on P_s. That is, $P_s \subseteq X$. Similarly choose a minimal certification $\mathcal{A}'_s \subseteq \mathcal{A}'$ of f on P_s. The same lemma says that f agrees with an element of $\mathcal{A}'_s$ everywhere on P_s.

If two polytopes of the form P_s overlap, their intersection appears as a face of each and is itself of the form P_s. Let $\mathcal{P}$ be the polytopes of form P_s not contained in any other. For each $P \in \mathcal{P}$ chose a function A_P in $\mathcal{A}_f$ that agrees with f on P. We may now proceed with the construction just as before, starting with the definition of the functions denoted $f_{P,Q}$ (where we understand $\mathcal{D}_P$ to be $\mathcal{D}_s$ for $s \in \overset{\circ}{P_s}$).

We may put this material together to obtain:

21.3 THEOREM: *The full subcategory of finitely presented scales is dual to the category of closed piecewise $\mathbb{D}$-affine sets and* CPDA *maps.*

A quite different notation presents itself, one that emphasizes the pivotal role played by $\mathtt{I}$ (to use a popular phrase—avoided by those bothered by etymology—it is the duality's "schizophrenic object"). Tradition insists that we change notation. Instead of $\mathcal{CPDA}(X)$ we'll use $\boldsymbol{X}^\star$ The functor that sends X to $X^\star$ may be viewed as an contravariant algebra-valued representable functor, that is, we could also denote $X^\star$ as $(X, \mathtt{I})$ where $\mathtt{I}$ is viewed as an object in the category of closed piecewise $\mathbb{D}$-affine sets equipped therein with its structure as a scale (just as can any equational theory in any category with finite products) thereby endowing $(X, \mathtt{I})$ with a scale structure for any X and endowing $(f, \mathtt{I})$ with the status of scale map for any $f : X \to X'$

When S is a scale, let $\boldsymbol{S}^\star$ denote the space of $\mathtt{I}$-valued scale-homomorphisms on S. Because $\mathtt{I}$-valued scale maps are known by their kernels (and because all simple scales are uniquely embeddable in $\mathtt{I}$) $S^\star$ is naturally equivalent to $\mathrm{Max}(S)$

The fact that this pair of functors is an equivalence of categories is equivalent to the "adjunction maps" $X \to X^{\star\star}$ and $S \to S^{\star\star}$ being isomorphisms. To establish the first isomorphism suppose that X is a subset of $\mathtt{I}^n$ of the form $h^{-1}(\top)$ where h is continuous piecewise $\mathbb{D}$-affine. Then $X^\star$ may be taken as $F_n/((h))$ where F_n is the free scale on n generators, $\mathtt{I}\langle x_1, x_2, \ldots, x_n\rangle$. The adjunction map $X \to X^{\star\star}$ sends $x \in X$ to the evaluation map that sends $f \in X^\star$ to $f(x)$. The semi-simplicity of $F_n/((h))$ says that $X \to X^{\star\star}$ is monic. To prove that it's onto, let $g \in X^{\star\star}_,$ that is, $g : X^\star \to \mathtt{I}$. We know that there's $x \in \mathtt{I}^n$ such that $F_n \to X^\star \to \mathtt{I}$ is the evaluation map at x. We need only show that x is in X. Suppose not. It suffices to find $k \in X^\star_,$ that is, $k : X \to \mathtt{I}$ such that $g(k) \neq k(x)$. It's easy: take $k = h$. The argument for $S \to S^{\star\star}$ is essentially the same.

It is worth noting that the standard notion of homotopy translates rather nicely into this setting. The "co-cylinder" over $S^\star$ is $S[\mathrm{U}]^\star$ where $S[\mathrm{U}]$ is the "polynomial scale" over S, that is, terms in a fresh variable U built from elements in S. A pair of maps $f_\perp, f_\top : T \to S$ gives rise to a pair of "co-homotopic" maps from $S^\star$ to $T^\star$ iff there is a map $H : T \to S[\mathrm{U}]$ such that for $e = \perp, \top$ we have $T \overset{H}{\to} S[\mathrm{U}] \overset{v_e}{\to} S \;=\; T \overset{f_e}{\to} S$ where v_e is the map that evaluates a polynomial at e. Brouwer's simplicial approximation theorem [101] is just what is needed for:

21.4 THEOREM: *The homotopy category of continuous maps between finitely triangulable spaces is dual to an algebraically defined quotient category of the category of f.p.scales.*

A closing comment: the transitivity of homotopy uses—directly—the coalgebra structure. If $f_\perp$ is co-homotopic to $f_\odot$ via $H_\perp$ and $f_\odot$ to $f_\top$ via $H_\top$ then we need to put $H_\perp$ and $H_\top$ together to create $H_\odot$ such that for $e = \perp, \top$

$$T \xrightarrow{H_\odot} S[\mathrm{U}] \xrightarrow{v_e} S \;=\; T \xrightarrow{f_e} S.$$

Using the comment on p20, $H_\odot$ may be constructed as:

$$f_\odot \lhd \left(\left(v_{\underset{\mathrm{U}}{\vee}}(H_\perp)\right) \middle| \left(v_{\underset{\mathrm{U}}{\wedge}}(H_\top)\right)\right). \qquad [102]$$

22. Complete Scales

Consider the smallest full subcategory of scales that includes $\mathtt{I}$ and is closed under the formation of limits. The construction of left-limit–closures of subcategories can be complicated, but the injectivity of $\mathtt{I}$ in the category of scales makes the job easier: it is the full subcategory of scales that appear as equalizers of pairs of maps between powers of $\mathtt{I}$. That is, first take the full subcategory of objects of the form $\prod \mathtt{I}$ and then add all equalizers of pairs of maps between them. It is clear that such are closed under the formation of products. To see that the full subcategory of such is closed under equalizers let X be the equalizer of a pair of maps from $\prod_I \mathtt{I}$ to $\prod_J \mathtt{I}$ and Y the equalizer of a pair of maps from $\prod_K \mathtt{I}$ to $\prod_L \mathtt{I}$. Let f, g be a pair of maps from X to Y. The injectivity of $\mathtt{I}$ allows us to extend f and g to maps from $\prod_I \mathtt{I}$ to $\prod_K \mathtt{I}$. The equalizer of f, g is constructable as the equalizer of the resulting pair of maps from $\prod_I \mathtt{I}$ to $\prod_J \mathtt{I} \times \prod_K \mathtt{I}$.

Any locally small category constructed as the left-limit–closure of a single object is automatically a reflective subcategory. We will call its objects **complete scales**. They have a number of alternative characterizations.

Say that a map of scales, $A \to B$ is a **weak equivalence** (a phrase borrowed from homotopy theory) if it is carried to an isomorphism by the set-valued functor $(-, \mathtt{I})$, or—put another way—if every $\mathtt{I}$-valued map from A factors uniquely through $A \to B$. A scale S is complete iff $(-, S)$ carries all weak equivalences to isomorphisms. The "lluf subcategory" (that is, one that contains all objects) of equivalences falls apart into connected components, one for each isomorphism type of complete scales. They are precisely the objects that appear as weak terminators in their components.

The most algebraic description of the category of complete scales is as a category of fractions, to wit, the result of formally inverting all the weak equivalences. (All full reflective subcategories are describable: they are always equivalent to the result of formally inverting all the maps carried to isomorphisms by the reflector functor.)

A quite different description of complete scales—one that appears not to be algebraic—is in terms of a metric structure. In this setting it is useful to take $\mathtt{I}$ to be the unit interval $[0, 1]$ and $\mathbb{I}$ the dyadic rationals therein. The **intrinsic pseudometric** (of diameter one) on a

[101] Everyone seems to agree that Brouwer intended, with this theorem, to transform topology into something we could get our hands on. Or, as we would say it now, something worthy of the name "algebraic topology."

[102] Yes—to my amazement—this is a scale homomorphism. If $t \in T$ then $H_\odot(t)$ is, for each U, either $f_\perp(t)$ or $f_\top(t)$. Hence for each U it is equal to a scale homomorphism.

scale S is most easily defined—in the presence of the axiom of choice—by taking the distance from x to y as $\sup_{f:S\to \mathrm{I}} |f(x) - f(y)|$. Such is a metric (not just a pseudometric) iff S is semi-simple. It is, further, a complete metric iff S is a complete scale (indeed, the reflection of an arbitrary scale into the subcategory of complete scales may be described—metrically—as the usual metric completion of the scale viewed as a pseudometric space).

The intrinsic metric may be defined directly without recourse to the axiom of choice. Define the **intrinsic norm** of $x \in S$ as $||x|| = \inf \{ q \in \mathbb{I} : x \leqq q \}$ and the distance between x and y as $||x \,\dot{\multimap\!\!\!\circ}\, y||$ (recall that $\dot{\multimap\!\!\!\circ}$ is the dotting operation applied to $\circ\!\!-\!\!\circ$). (It is easy to verify that on the unit interval $x \,\dot{\multimap\!\!\!\circ}\, y = |x{-}y|$.) To see that this definition agrees with the previous in the presence of the axiom of choice we need to show that $||x|| = \sup_{f:S\to \mathrm{I}} f(x)$. Clearly $f(x) \leqq ||x||$ for all $f : S \to \mathrm{I}$. If $||x|| = 0$ we are done. For the reverse semiquation when $||x|| > 0$ it suffices to find a single $f : S \to \mathrm{I}$ such that $f(x) = ||x||$ and for that it suffices—in the presence of the axiom of choice—to find a proper $\top$-face that contains $\{q_n \multimap x\}_n$ for a strictly ascending sequence $\{q_n\}_n$ of dyadic rationals approaching $||x||$. The $\top$-face generated by $\{q_n \multimap x\}_n$ is the ascending union of the principal $\top$-faces $\{ ((q_n \multimap x)) \}_n$ hence it suffices to show that each $((q_n \multimap x))$ is proper. It more than suffices to find a linear quotient in which $q_n \leqq x$. The linear representation theorem says that if there were no such linear quotient then $x < q_n$, directly disallowed by the choice of the q_ns.

23. Scales vs. Spaces

The main goal of this section is to show that the category of compact-Hausdorff spaces is dual to the category of complete scales. Using the notation already introduced, the equivalence functor from compact Hausdorff spaces to complete scales sends X to $\mathcal{C}(X)$, the scale of continuous I-valued maps on X. The equivalence functor from complete scales to compact Hausdorff spaces sends S to $\mathrm{Max}(S)$, the set of maximal $\top$-faces on S, topologized by the standard "hull-kernel" topology.

As with the duality between f.p.scales and closed piecewise $\mathbb{D}$-affine sets, tradition insists that we change notation. Instead of $\mathcal{C}(X)$ we use $X^\star$. Again, the functor that sends X to $X^\star$ may be viewed as an contravariant algebra-valued representable functor, that is, we could also denote $X^\star$ as (X, I) where I is viewed as a scale algebra in the category of topological spaces. It is clear from the metric characterization that $X^\star$ is a complete scale.

If S is a scale let $S^\star$ denote the set of I-valued scale-homomorphisms on S, topologized by taking as a basis all sets of the form, one for each $s \in S$:

$$U_s \;=\; \{ f \in S^\star : f(s) < \top \}$$

The fact that $\mathrm{Max}(S)$ and $S^\star$ describe the same space rests on the fact that I-valued scale-homomorphisms are known by their kernels. (The fact that the hull-kernel topology describes the same space is easily verified: U_s corresponds to the complement of the hull of the principal $\top$-face $((s))$.)

We will find useful the formulas:

$$\begin{aligned} U_s \cap U_t &= U_{s\vee t} \\ U_s \cup U_t &= U_{s\wedge t} \\ U_s &= U_{\hat{s}} \\ U_\top &= \emptyset \\ U_\bot &= S^\star \end{aligned}$$

23.1 LEMMA: *Spaces of the form $S^\star$ are compact-Hausdorff.*

For the Hausdorff property let f, g be distinct elements of $S^\star$ and chose $a \in S$ such that $f(a) \neq g(a)$. We may assume without lose of generality that $f(a) < g(a)$. Let $q \in \mathbb{I}$ be such that $f(a) < q < g(a)$. Then $f \in U_{q \multimap a}$ and $g \in U_{a \multimap q}$.[103] The equation of linearity yields $U_{q \multimap a} \cap U_{a \multimap q} = U_{(q \multimap a)\vee(a \multimap q)} = U_\top = \emptyset$. For compactness let S' be a subset of S. The necessary and sufficient condition that the family of sets $\{ U_s : s \in S' \}$ be a cover of $S^\star$ is that the $\top$-face generated by S' is entire (because the elements $f \in S^\star$ not in any U_s are precisely those such that $S' \subseteq \ker(f)$). A $\top$-face is entire iff it contains $\bot$. Hence there must be $s_1, s_2, \ldots, s_n \in S'$ such that $\bot$ is the result of applying $\top$-zooming a finite number of times to the element $s_1 \wedge s_2 \wedge \ldots \wedge s_n$. And that is enough to tell us that $U_{s_1} \cup U_{s_2} \cup \cdots \cup U_{s_n} = U_{s_1 \wedge s_2 \wedge \cdots \wedge s_n} = U_\bot = S^\star$

For each $s \in S$ we obtain the "evaluation map" from $S^\star$ to $\mathbf{I}$ that sends $f \in S^\star$ to $f(s) \in \mathbf{I}$. Yet another description of the topology on $S^\star$ is as the weakest topology that makes all these evaluation maps continuous: given $q < r \in \mathbb{I}$ the inverse image of the open interval $(q, r) \subseteq \mathbf{I}$ is $U_{(s \multimap q)\vee(r \multimap s)}$ (and, of course U_s is the inverse image of $\mathbf{I}\backslash\{\top\}$).

For an arbitrary space X and element $x \in X$ the evaluation map in $X^{\star\star}$ that sends $f \in X^\star$ to $f(x) \in \mathbf{I}$ is clearly a homomorphism. The natural map $X \to X^{\star\star}$ that sends each point in X to its corresponding evaluation map is continuous (the inverse image of $U_f \subseteq X^{\star\star}$ is the f-inverse-image in X of the open subset $\mathbf{I}\backslash\{\top\}$.)

23.2 THEOREM: *If X is compact-Hausdorff then $X \to X^{\star\star}$ is a homeomorphism*

It is monic because of the Urysohn lemma. To see that it is onto, let $H : X^\star \to \mathbf{I}$ be an arbitrary scale-homomorphism. Because $\ker(H)$ is maximal it suffices to find x such that $\ker(H)$ is contained in the kernel of the evaluation map corresponding to x, that is, it suffices to find x such that $H(f) = \top$ implies $f(x) = \top$, or put another way, to find x in $\bigcap_{f\in\ker(H)} f^{-1}(\top)$. First note that if $f^{-1}(\top) = \emptyset$ then there is $q \in \mathbb{I}$ such that $f < q < \top$, hence $H(f) < q$ forcing $f \notin \ker(H)$. The compactness of X therefore says that it suffices to show that the finite-intersection property holds for the family of closed sets $\{ f^{-1}(\top) : f \in \ker(H) \}$. But this family is closed under finite intersection: $f^{-1}(\top) \wedge g^{-1}(\top) = (f \wedge g)^{-1}(\top)$ and $\ker(H)$ is clearly closed under finite intersection.

The natural map $S \to S^{\star\star}$ that sends each element in S to its corresponding evaluation map is a homomorphism.

23.3 THEOREM: *If S is a complete scale then $S \to S^{\star\star}$ is an isomorphism.*

The proof is an immediate consequence of

23.4 THEOREM: *If X is compact-Hausdorff then the necessary and sufficient condition for a subscale, S, of $X^\star$ to be dense (under the intrinsic metric) is that S separates the points of X, that is, for every two points $x, y \in X$ there exists $f \in S$ such that $f(x) \neq f(y)$.*

Note that the necessity uses the Urysohn lemma. The proof is much easier than its model, the Stone-Weierstrass theorem (or is it—in this case—Stone-without-Weierstrass?). We first establish that S has the "two-point approximation property," that is, for every pair $a, b \in \mathbb{I}$ and every pair of distinct points $x, y \in X$ there is $f \in S$ with $f(x) = a$ and $f(y) = b$. If $a = b$ we can, of course, take f to be that constant. Otherwise we can assume without loss of generality that $a < b$. Start with any f such that $f(x) \neq f(y)$, as insured by the hypothesis. If $f(x) > f(y)$ replace f with $\dot{f}$. Let $c \in \mathbb{I}$ be such that $f(x) < c < f(y)$. There exists n such that $(c\triangleleft)^n f(x) = \bot$ and $(c\triangleleft)^n f(y) = \top$. Replace f with $(c\triangleleft)^n f$ to achieve $f(x) = \bot$ and

[103] There is actually a canonical choice: take q to be of minimal denominator (it's unique).

$f(y) = \top$. Finally, replace that f with $a \vee (f \wedge b)$ to achieve $f(x) = a$ and $f(y) = b$.

We can now repeat the Stone argument. Let X be a compact space and S a sublattice in $X^\star$ with the two-point approximation property. Given any $h \in X^\star$ and $\varepsilon > 0$ we wish to find $f \in S$ such that the values of f and h are everywhere within ε of each other. For each pair of points $x, y \in X$ let $f_{x,y} \in S$ be such that $f_{x,y}(x)$ is within ε of $h(x)$ and $f_{x,y}(y)$ is within ε of $h(y)$. Define the open set

$$U_{x,y} \quad = \quad \{\, z \in X \,:\, f_{x,y}(z) < h(z) + \varepsilon \,\}.$$

It is best to regard $\mathbf{I}$ here as a fixed closed interval in $\mathbb{R}$. Since $y \in U_{x,y}$ we know that for fixed x the family $\{U_{x,y}\}_y$ is an open cover. Let $y_1, y_2, \ldots, y_m$ be such that $U_{x,y_1} \cup U_{x,y_2} \cup \cdots \cup U_{x,y_m} = X$ and define $f_x = f_{x,y_1} \wedge f_{x,y_2} \wedge \cdots \wedge f_{x,y_m}$. Then $f_x(x) > h(x) - \varepsilon$ and for all z we have $f_x(z) < h(x) + \varepsilon$ Now for each x define the open set

$$U_x \quad = \quad \{\, z \in X \,:\, f_x(z) > h(z) - \varepsilon \,\}.$$

Since $x \in U_x$ we know that the family $\{U_x\}_x$ is an open cover. Let $x_1, x_2, \ldots x_n$ be such that $U_{x_1} \cup U_{x_2} \cup \cdots \cup U_{x_n} = X$. Finally define $f = f_{x_1} \vee f_{x_2} \vee \cdots \vee f_{x_n}$. [104]

23.5 THEOREM: *The full subcategory of complete scales is dual to the category of compact-Hausdorff spaces.*

For a remarkably algebraic definition of the category of compact Hausdorff spaces: start with the category of scales; chose an object that has a unique endomorphism and is a weak terminator for all objects other than the terminator; call that object $\mathbf{I}$; formally invert all maps that the contravariant functor $(-, \mathbf{I})$ carries to isomorphisms; take the opposite category. One can replace $\mathbf{I}$ with any non-trivial injective object (as will be seen in the next section all injective scales are retracts of cartesian powers of $\mathbf{I}$). Another choice is first to restrict to the full subcategory of semi-simple scales and then formally invert the mono-epis (obtaining what some might call its "balanced reflection").[105]

24. Injective Scales, or: Order-Complete Scales

As complete scales are to the study of continuous maps we expect that injective scales will be to measurable functions. As we will see, all injective scales come equipped with a (necessarily unique) chromatic structure but—fortunately for that expectation—maps between them need not preserve that structure. Let us lay out the groundwork.

If an object (in any category) is injective, then it is clearly an **absolute retract**, that is, whenever it appears as a subobject it appears as a retract. The converse need not hold, even for models of an equational theory.[106] As we will see though, absolute retracts in the category of scales are, indeed, injective.

[104] Yes, the argument of this paragraph fails if X has only one element.

[105] In the full subcategory of semi-simples a map is an epi iff its image is an order-dense subset of the target (and—as always for models of equational theories—monos are one-to-one). But for arbitrary scales the characterization of epimorphisms is more complicated. If we work with $\mathbb{Q}$-scales, that is, scales for which multiplication by each rational in the standard interval is defined, then the only epis are onto (hence all mono-epis are isos). For a proof, it suffices (as usual) to consider "dense subobjects," that is, those whose inclusion maps are epi. Suppose $S' \subseteq S$ is dense. Let V be the enveloping $\mathbb{Q}$-vector space of S (defined, of course, analogously to the enveloping $\mathbb{D}$-module) with its inherited partial ordering. For any $s_0 \in S \setminus S'$ use the axiom of choice to find a map $f : V \to \mathbb{Q}$ with S' in its kernel such that $f(s_0) \neq 0$. Partially order $V \oplus \mathbb{Q}$ lexicographically and define a new scale T as the interval from $\langle \bot, 0 \rangle$ to $\langle \top, 0 \rangle$ with its induced scale structure. The two scale homomorphisms that send $s \in S$ to $\langle s, 0 \rangle$, in the first case, and $\langle s, f(s) \rangle$, in the second, agree on S' and disagree on s_0. For the general scale case, the condition for S' to be dense in S is that it be a "pure" subscale: $ns \in S'$ implies $s \in S'$ for all integers $n > 0$.

[106] The Sierpiński monoid ($\{0,1\}$ under multiplication) is an absolute retract in the category of commutative monoids: whenever it is a submonoid the "least map" (least with respect to the order induced from $\{0,1\}$) from the ambient monoid to the Sierpiński monoid (to wit, the characteristic map of the subgroup of units) is a retraction. But the least map from the one-generator monoid (the natural numbers) does not extend to a map from the one-generator group (the integers). Curiously

24.1 LEMMA: *A scale that is an absolute retract is order-complete (that is, it is not just a lattice but a complete lattice).*

Let E be an absolute retract and $L \subseteq E$ an arbitrary subset, $U \subseteq E$ the set of upper bounds of L. Freely adjoin an element b to E to obtain the "polynomial scale" $E[b]$ and let $\mathcal{F}$ be the $\top$-face generated by the elements $\{\ \ell \multimap b : \ell \in L\ \}$ and $\{\ b \multimap u : u \in U\ \}$. Once we know that $E \to E[b]/\mathcal{F}$ is an embedding we are done because then a retraction $f : E[b]/\mathcal{F} \to E$ necessarily sends b to a least upper bound of L. The fact that $E \to E[b]/\mathcal{F}$ is an embedding is equivalent to the disjointness of $\mathcal{F}_\star$ and $E_\star$ (recall that the "lower star" removes the top). If $x \in E_\star$ were in $\mathcal{F}$ then a finite number of the generators of $\mathcal{F}$ would account for it. But given $\ell_1 \multimap b, \ell_2 \multimap b, \ldots, \ell_m \multimap b$ and $b \multimap u_1, b \multimap u_2, \ldots, b \multimap u_n$ we may easily obtain a retraction of $E[b]$ back to E by sending b to $u_1 \wedge u_2 \wedge \cdots \wedge u_n$. The kernel of any retraction of E is, of course, disjoint from $E_\star$. But this kernel contains the listed finite number of elements (indeed it contains $\ell \multimap b$ for all $\ell \in L$).

24.2 COROLLARY: *Absolute retracts are chromatic scales.*

Easily enough: $\overline{x} = \bigvee_m (\bot\triangleleft)^m x$ and, hence, $\underline{x} = \bigwedge_m (\top\triangleleft)^m x$.

24.3 LEMMA: *Absolute retracts are semi-simple scales.*

Semi-simplicity is equivalent to $\top$ being the least upper bound of $\mathbb{I}_\star$. But a least upper bound of $\mathbb{I}_\star$ is necessarily invariant under $\top$-zooming. There is only one such element larger than $\odot$ (check in any linear scale) and that element is $\top$.

24.4 COROLLARY: *Absolute retracts are injective scales.*

An absolute retract, being semi-simple, can be embedded in a cartesian power of the injective object $\mathbf{I}$. [107] A cartesian power of an injective is injective, that is, every absolute retract can be embedded in an injective which, of course, is the necessary and sufficient condition for an absolute retract to be injective (any retract of an injective is injective).

As is the case for any equational theory, a model is an absolute retract iff it has no proper **essential extensions**. We recall the definitions: a monic $A \to B$ is essential if whenever $A \to B \to C$ is monic it is the case that $B \to C$ is monic. For models of an equational theory this translates to the condition that every non-trivial congruence on B remains non-trivial when restricted to A. Note that for any monic $A \to B$ we may use Zorn's lemma to obtain a congruence maximal among those that restrict to the trivial congruence on A thus obtaining a map $B \to C$ such that $A \to B \to C$ is an essential extension. It follows that if A has no essential extensions other than isomorphisms then A is an absolute retract. (The converse is immediate.)

24.5 LEMMA: *A scale B is an essential extension of a subscale $A \subseteq B$ iff $A_\star$ is co-final in $B_\star$*

Because essentiality is clearly equivalent to every non-trivial $\top$-face in B meeting A non-trivially; and clearly a principal $\top$-face, $(\!(b)\!) \subseteq B$ meets A non-trivially iff there is

the Sierpiński monoid is injective in the full subcategory of finite commutative monoids, indeed, in the full subcategory of locally finite commutative monoids (which in the commutative case is the same as saying each one-generator submonoid is finite). The full subcategory of idempotent commutative monoids—semi-lattices as they are usually called—is both a reflective and co-reflective subcategory of the category of all commutative monoids. If one restricts to locally finite commutative monoids then not only is it both reflective and co-reflective but the same functor delivers both the reflection and co-reflection: the reflection map sends each element of a locally finite commutative monoid to the unique idempotent element that appears in the sequence of its positive powers. Since the functor also delivers co-reflections it preserves monomorphisms and hence the injectivity of the Sierpiński monoid follows from its injectivity among semi-lattices. If one deems them meet-semi-lattices then maps into the two-element meet-semi-lattice are the characteristic maps of filters and it is easy to see that any such characteristic map extends to any larger meet-semi-lattice. (Note that—quite unusual for injective objects—the axiom of choice is not used in extending maps to the Sierpiński monoid.)

[107] For the injectivity of $\mathbf{I}$ see Theorem 10.7 (p29).

$a \in A_\star$ such that $b < a$. The converse for Lemma 24.1 (p54):

24.6 LEMMA: *If a scale is order-complete then it is an absolute retract..*

First a lemma: if $A_\star$ is co-final in $B_\star$ then for any $a \in A$ and $b \in B$ such that $b < a$ there exists $a' \in A$ with $b < a' < a$ because we may use the order-isomorphism $a \multimap (-)$ from $[b, a]$ to $[a \multimap b, \top]$ [108] to obtain $a'' \in A$ such that $a \multimap b < a'' < \top$. The inverse isomorphism thus delivers an element in A (to wit, $a' = a \overset{\wedge}{|} a''$) strictly between b and a. The dual lemma: if $a < b$ we may find $a' \in A$ such that $a < a' < b$ (simply apply the previous case to $\dot{b} < \dot{a}$).

Consider an order-complete scale A and essential extension $A \subseteq B$. Given $b \in B$ let $a \in A$ be the greatest lower bound of the A-elements above b. Using the dual lemma we reach a contradiction from the strict semiquation $a < a \vee b$ (because if $a' \in A$ were such that $a < a' < a \vee b$ then a would not be the *greatest* lower bound). Hence $a = a \vee b$, that is, $b \leqq a$. Using the lemma (as opposed to its dual) we reach a contradiction from the strict semiquation $b < a$ (because if $a' \in A$ were such that $b < a' < a$ then a would not be a lower bound of the A-elements above b). Thus $a = b$. That is, every element in B is in A.

In great generality—in particular for the models of any equational theory—a maximal essential extension of an object is injective, indeed, it is minimal among injective objects in which the object can be embedded. Such maximal-essential/minimal-injective extensions are called **injective envelopes**.[109] (If an object can be embedded in an injective then each of its essential extensions must appear therein and we may find one that is maximal.)

For scales we have noted that $A \subseteq B$ is essential iff $A_\star$ is cofinal in $B_\star$. Semi-simple scales, we know, can be embedded in injectives (to wit, cartesian powers of $\mathbf{I}$) hence have injective envelopes (which are unique up to—perhaps many—isomorphisms).[110]

24.7 THEOREM: *Every injective scale is the injective envelope of its subscale generated by its extreme points.*

Given an injective scale E and an element $a < \top$ we need $n \in \mathbb{N}$ and an extreme point e such that $a \leqq (\top|)^n e < \top$. Using the semi-simplicity of E we can specialize to the case that E is a subscale of a cartesian power $\prod_K \mathbf{I}$. [111] Take n to be such that $(\top|)^n \bot$ is *not* a lower bound of a (if there were no such n then a would be $\top$). We first describe an element $e \in \prod_K \mathbf{I}$ by stipulating its value for each co-ordinate $i \in K$:

$$e_i \;=\; \top \;\texttt{if}\; (\top|)^n \bot \leqq a_i \;\texttt{else}\; \bot$$

e satisfies the three equations:

$$\top \;=\; e \vee \dot{e} \;=\; a \multimap \big((\top|)^n e\big) \;=\; e \multimap \big((\top|)^n \bot \multimap a\big)$$

The 1st equation says that e is an extreme point. The 2nd equation says that $a \leqq (\top|)^n e$. The 3rd equation ensures that $e < \top$, hence $(\top|)^n e < \top$. Now use injectivity to obtain an element in E satisfying the same three equations.

24.8 THEOREM: *Order-complete/injective scales are precisely those scales of the form $\mathcal{C}(X)$ where X is an extremely disconnected compact Hausdorff space (that is, one in which the closure of every open set is open or—as sometimes called— "clopen").*

---

[108] This was discussed at the end of Section 4 (p14–17).

[109] Sometimes "injective hulls,"

[110] If it is not semi-simple it will have essential extensions of unbounded cardinality, indeed, in Section 12 (p31–32) we saw this phenomenon for the Richter scale when we saw, first, that is appears co-finally in every non–semi-simple SDI and, secondly, that there are non–semi-simple SDIs of unbounded cardinality.

[111] Yes, of course we could take $K = \mathrm{Max}(E)$ but we needn't since its topology has no role in this proof.

Note first that any retract of a (metrically) complete scale is complete, hence injective scales, being retracts of cartesian powers of $\mathbb{I}$, are necessarily of the form $\mathcal{C}(X)$. We need to show that the order-completeness implies that X is extremely disconnected. Let $U \subseteq X$ be open. Define $\mathcal{F} \subseteq \mathcal{C}(X)$ to be the set of all continuous functions from X to $\mathbb{I}$ that are constantly equal to $\bot$ on the complement of U and let $\mathcal{U} \subseteq \mathcal{C}(X)$ be the set of upper bounds of $\mathcal{F}$ and let $\mathcal{L}$ be the set of lower bounds of $\mathcal{U}$. The order-completeness of $\mathcal{C}(X)$ says that there is a (necessarily unique) continuous $g \in \mathcal{L} \cap \mathcal{U}$. The Urysohn lemma says that for every $x \in U$ there is an element in $\mathcal{F}$ that sends x to $\top$ hence every element in $\mathcal{U}$ sends all of U to $\top$. The Urysohn lemma also says that for every point x in the complement of $\overline{U}$ there is a function in $\mathcal{U}$ that sends x to $\bot$ hence we know that g is constantly equal to $\top$ on U, therefore, on its closure $\overline{U}$. And, of course, any function equal to $\top$ everywhere on U is is an upper bound of $\mathcal{L}$

Dually, we know that for every $x \notin \overline{U}$, the closure of U, there is an upper bound of $\mathcal{L}$ that sends x to $\bot$ hence the least upper bound of $\mathcal{L}$ must be constantly equal to $\top$ on $\overline{U}$ and $\bot$ on its complement. It is the (continuous!) characteristic function of $\overline{U}$ which means, of course, that $\overline{U}$ is clopen.

The proof of the converse makes use of an unexpected subject. Let $\mathring{\mathbb{I}}$ denote the initial scale with the endpoints removed and P a complete poset. Define $\mathcal{Q}(P)$ to be the set of sup-preserving functions from $\mathring{\mathbb{I}}$ to P. We will use the fact that $\mathcal{Q}(P)$ is also a complete lattice. And that use will take advantage of the Urysohn method for obtaining continuous $\mathbb{I}$-valued functions.[112]

Given a extremely disconnected space X take P to be the boolean algebra of clopens. Given $U \in \mathcal{Q}(P)$ define $f : X \to \mathbb{I}$ by $f(x) = \inf\{\, t : x \in U_t \,\}$ and obtain that for all $s \in \mathbb{I}$ that $f^{-1}[\bot, s) = \bigcup_{r<s} U_r$ is necessarily open and that $f^{-1}[\bot, s] = \bigcap_{t>s} U_t$ is necessarily closed which, note, establishes the continuity of f. Moreover every $f \in \mathcal{C}(X)$ is obtainable from a unique element in $\mathcal{Q}(P)$, to wit, the element defined by taking U_s to be the closure of $f^{-1}[\bot, r)$. All of which shows that $\mathcal{C}(X)$ is order-complete. (The correspondence between $\mathcal{C}(X)$ and $\mathcal{Q}(P)$ is contravariant.)[113]

[In the TAC version it was erroneously stated that "Order-complete scales are precisely those (metrically) complete scales that are chromatic." For a complete chromatic scale that's not order-complete see [153], p82 (where it appears as a footnote for a proof that the converse implication—that order-complete/injective scales are chromatic—is correct).]

Since we have identified the injective objects in the full category of complete scales it follows that we have identified the projectives in its dual category, the category of compact Hausdorff spaces.[114] Given such a space X let Y be the set of ultrafilters on the (discrete) set X. Y is, of course, the "compactification" of that discrete set, that is, the reflection of the discrete space in the full subcategory of compact Hausdorff spaces. The canonical map from Y back to X is a retraction of X and, hence, $\mathcal{C}(X)$ is a retract of $\mathcal{C}(Y) = \prod_X \mathbb{I}$ which, being a cartesian product of injective objects, is, of course, injective.

In the category of compact Hausdorff spaces Gleason constructed the minimal projective cover of a space X as the Stone space of the Boolean algebra of regular closed sets of X (those

[112] Urysohn showed that if P is the lattice of open sets of a space X and $U \in \mathcal{Q}(P)$ has the "Urysohn property," to wit, the property that the closure of U_s is contained in U_t for all $s < t$ then we obtain continuous $f : X \to \mathbb{I}$ such that $f(x) = \inf\{\, t : x \in U_t \,\}$. Moreover given any continuous f we obtain a Urysohn element by $U_s = f^{-1}[\bot, s)$.

[113] It's worth noting just where this argument fails when P is taken as the lattice of open subsets: we lose the Urysohn property when we take infinite joins.

[114] First done by Andrew Gleason **Projective topological spaces.** *Illinois J. Math.* 2 (1958), p482–489. But this result is an easy consequence of the fact that the injective objects in the category of boolean algebras are the complete boolean algebras: Sikorski, Roman **A theorem on extension of homomorphisms.** *Ann. Soc. Polon. Math.* 21 (1948), p332–335.

closed sets that are the closures of their interiors). That Stone space may be constructed as the set of ultrafilters of regular closed sets. The covering map is clear: send such an ultrafilter to the unique point in its intersection. It may be described also as the scale of "adjoint pairs" of semicontinuous $\mathbb{I}$-valued functions on X, that is pairs consisting of a lower- and an upper semicontinuous function where the lower semicontinuous function is the largest lower semicontinuous function less than the upper semicontinuous function, and dually.

25. *Diversion:* Finitely Presented Chromatic Scales

If we move to chromatic scales we remove the word "continuous" to obtain "piecewise $\mathbb{D}$-affine map" and the word "closed" to obtain "piecewise $\mathbb{D}$-affine set." The definitions are no longer as simple (we can not, for example, get by just with the existence of a $\mathbb{D}$-affine certification). The proofs of the parallel theorems, however, are easier. The category of finitely presented chromatic scales is dual to the category of piecewise $\mathbb{D}$-affine maps between piecewise $\mathbb{D}$-affine sets. Any piecewise $\mathbb{D}$-affine set is a disjoint union of "boundaryless simplices," to wit, those that result when all boundary points are removed from ordinary closed simplices (note that the 0-dimensional boundaryless simplex is not empty but a single point); an n dimensional piecewise $\mathbb{D}$-affine set may be described with an $(n+1)$-tuple of natural numbers, $\langle s_0, s_2, \ldots, s_n\rangle$ specifying the number of boundaryless simplices of each dimension (necessarily $s_n > 0$). Define the Euler characteristic, χ, as $s_0 - s_1 + \cdots + (-1)^n s_n$. Non-empty spaces turn out to be determined up to isomorphism by just two invariants: dimension and Euler characteristic.[115]

The chromatic scale corresponding to an n-dimensional space can not be generated with fewer than n generators. If $\chi = 1$ the corresponding chromatic scale may be taken to be the free chromatic scale on n generators, $\mathbb{I}[x_1, x_2, \ldots, x_n]$.

If $\chi > 1$ then the corresponding scale may be constructed as $\mathbb{I}[x_1, x_2, \ldots, x_n]/(\!(\underline{t})\!)$ where $t = (x_1 \multimap q_1) \vee (x_1 \circ\!\!-\!\!\circ q_2) \vee \cdots \vee (x_1 \circ\!\!-\!\!\circ q_\chi)$, one $\multimap$, the rest $\circ\!\!-\!\!\circ$s, and $\bot < q_1 < q_2 < \cdots < q_\chi = \top$ are $\mathbb{I}$-elements. As always for finitely presented chromatic scales this is isomorphic to an interval in the free algebra, to wit, $[\bot, \underline{t}\,]$. The free algebra splits as the product $[\bot, \underline{t}\,] \times [\,\underline{t}, \top]$ and we may dispatch the case of negative characteristic with the observation that the second factor is of the same dimension and the characteristics of the two factors add to one. For $\chi = 0$ we can use $\mathbb{I}[x_1, x_2, \ldots, x_n]/(\!(\overline{x_1})\!)$. Finally, when $n = 0$ the only corresponding chromatic scales are the cartesian powers of $\mathbb{I}$ (bear in mind that χ is necessarily non-negative and that the empty product is the one-element terminal scale).

The only time that we need more generators than dimension is when $n = 0$ and $\chi > 1$. (The corresponding chromatic scale is constructable with one generator: $\mathbb{I}[x]/(\!(\underline{v})\!)$ where $v = (x \circ\!\!-\!\!\circ q_1) \vee (x \circ\!\!-\!\!\circ q_2) \vee \cdots \vee (x \circ\!\!-\!\!\circ q_\chi)$ and $q_1 < q_2 < \cdots < q_\chi$.)

---

[115] Any such set is piecewise $\mathbb{D}$-affine isomorphic to one described by an $(n+1)$-tuple where $s_0 = s_1 = \cdots = s_{n-2} = 0$ and either $(s_{n-1} = 0)$ & $(s_n > 0)$ or $(s_{n-1} > 0)$ & $(s_n = 1)$ (the first possibility occurs precisely when $\chi \neq 0$ with signature $(-1)^n$). First, any k-dimensional boundaryless simplex with $k > 0$ is the disjoint union of one $(k-1)$-dimensional and two k-dimensional boundaryless simplices, which means that we may, without changing isomorphism type, increment by 1 any two adjacent s_is provided the right-hand one is already positive. A sequence of such increments—working from the top down—can guarantee that all the s_is are positive, indeed as big as we want them. We can then minimize the number of positive s_is by successively performing the reverse of such increments—working from the bottom up—until $s_0 = s_2 = \cdots = s_{n-2} = 0$. We then perform as many such "reverse increments" as we can on the pair s_{n-1}, s_n. The result is as advertised.

26. *Appendix:* Lattice-Ordered Abelian Groups

By a **lattice-ordered abelian group**, or **LOAG** for short, is meant, of course, an object with both an abelian-group and a lattice structure in which the lattice ordering is preserved by addition (that is, $x + (y \diamond z) = (x + y) \diamond (x + z)$ for either lattice-operation $\diamond$).[116]

There are many similarities and differences between the theories of LOAGs and of scales.

Among the similarities (each of which, I trust, is to be found somewhere in the literature):

▪ *The theory of* LOAG*s is a complete equational theory,* that is, every equation on its operators is either inconsistent or a consequence of its axioms (as exemplified in the last footnote). Every consistent equation holds for the LOAG of integers, $\mathbb{Z}$, because every consistent equation has a non-trivial model, every non-trivial model has a positive element (*e.g.,* $(x \vee 0) + ((-x) \vee 0)$ for any $x \neq 0$) and any positive element generates a sub-LOAG isomorphic to $\mathbb{Z}$, all of which says that the maximal consistent equational extension of the theory of LOAGs is—precisely—the theory of $\mathbb{Z}$. To verify that the equations in hand already provide that maximal consistent extension it thus suffices to show that every equation not a consequence of those axioms has a counterexample in $\mathbb{Z}$. It clearly suffices to find a counterexample in the rationals, $\mathbb{Q}$, because multiplying by a suitable positive integer would then yield a counterexample in $\mathbb{Z}$ and because the operations are continuous, it clearly suffices for that to find a counterexample in the reals, $\mathbb{R}$. The previous proof for the theory of scales can be easily replicated for this case. Or, if one wishes, we can reduce this case to that previous case. Given an equation with a counterexample in some LOAG we can first tensor with the dyadic rationals, $\mathbb{D}$, to obtain a $\mathbb{D}$-module and then chose an element we'll call $\top$ large enough so that the computation of the terms in the counterexample all lie in the interval $[-\top, \top]$. Replacing 0 with $\odot$, $-x$ with $\dot{x}$ and $x + y$ with $\odot\lhd (x|y)$ we obtain a scale with a counterexample for the given equation and from that we know that there is a counterexample in $\mathbb{R}$.

▪ *The free* LOAG *on n generators is the* LOAG *of continuous piecewise integral-linear* $\mathbb{R}$*-valued functions on* $\mathbb{R}^n$.[117] The functions in question are necessarily "radial:" $f(rx) = rf(x)$ for any $r > 0$, hence are determined by their values on the faces of the "standard cube" and that allows us to reduce to the result for free scales.

▪ *Every* LOAG *can be embedded in a product of linearly ordered abelian groups* (**TOAG**s). The

---

[116] We can simplify the definition by noting that we need only **truncation at zero**, $0 \vee x$, as a primitive. We will denote this truncation here as $\lfloor x \rfloor$. The two axioms: TRUNC-1: $x = \lfloor x \rfloor - \lfloor -x \rfloor$ TRUNC-2: $\lfloor x - \lfloor y \rfloor \rfloor = \lfloor \lfloor x \rfloor - \lfloor y \rfloor \rfloor$

Trunc-1 is justified by $x + (0 \vee -x) = (x + 0) \vee (x + (-x)) = x \vee 0$. For Trunc-2 it suffices to justify $\lfloor y \rfloor + \lfloor x - \lfloor y \rfloor \rfloor = \lfloor y \rfloor + \lfloor \lfloor x \rfloor - \lfloor y \rfloor \rfloor$. But $\lfloor y \rfloor + (0 \vee (x - \lfloor y \rfloor)) = (\lfloor y \rfloor + 0) \vee (\lfloor y \rfloor + (x - \lfloor y \rfloor)) = \lfloor y \rfloor \vee x = (y \vee 0) \vee x$ and $\lfloor y \rfloor + (0 \vee (\lfloor x \rfloor - \lfloor y \rfloor)) = (\lfloor y \rfloor + 0) \vee (\lfloor y \rfloor + (\lfloor x \rfloor - \lfloor y \rfloor) = \lfloor y \rfloor \vee \lfloor x \rfloor = (y \vee 0) \vee (x \vee 0)$. See Section 46 (p114–119) for subscorings.)

Use the truncation operator to define: $x \vee y = x + \lfloor y - x \rfloor$ The idempotence of $\vee$ is easily seen to be equivalent to what we'll call Trunc-0, $\lfloor 0 \rfloor = 0$, which can be proven by $\lfloor 0 \rfloor = \lfloor \lfloor 0 \rfloor \rfloor - \lfloor 0 - \lfloor 0 \rfloor \rfloor = \lfloor \lfloor 0 \rfloor \rfloor - \lfloor \lfloor 0 \rfloor - \lfloor 0 \rfloor \rfloor = \lfloor \lfloor 0 \rfloor \rfloor - \lfloor 0 \rfloor = \lfloor \lfloor \lfloor 0 \rfloor \rfloor - \lfloor 0 \rfloor \rfloor - \lfloor \lfloor 0 \rfloor - \lfloor \lfloor 0 \rfloor \rfloor \rfloor = \lfloor \lfloor 0 \rfloor - \lfloor 0 \rfloor \rfloor - \lfloor \lfloor \lfloor 0 \rfloor \rfloor - \lfloor \lfloor 0 \rfloor \rfloor \rfloor = \lfloor 0 \rfloor - \lfloor 0 \rfloor = 0$.

The fact that addition distributes with $\vee$ is immediate. For commutativity note that $x \vee y = y \vee x$ translates to $x + \lfloor y - x \rfloor = y + \lfloor x - y \rfloor$ and that rearranges to $x - y = \lfloor x - y \rfloor - \lfloor y - x \rfloor$, an instance of Trunc-1.

For associativity note that by adding $\lfloor y \rfloor$ to both sides of Trunc-2 we obtain $\lfloor y \rfloor \vee x = \lfloor y \rfloor \vee \lfloor x \rfloor$, making it clear that $\lfloor y \rfloor \vee x = \lfloor y \rfloor \vee \lfloor x \rfloor = y \vee \lfloor x \rfloor$ and hence $(y \vee 0) \vee x = \lfloor y \rfloor \vee x = y \vee \lfloor x \rfloor = y \vee (0 \vee x)$ which—together with distributivity with addition—easily yields full associativity.

(Trunc-2 is stronger than associativity—it has Trunc-0 built into it. It was chosen as an axiom not for its strength but for its simplicity: the truncation equation equivalent to associativity is $\lfloor x + \lfloor y - x \rfloor \rfloor = \lfloor x \rfloor + \lfloor y - \lfloor x \rfloor \rfloor$. For a separating example take the positive rationals under multiplication and define the associative "join" operation to be ordinary addition. The truncation operator is then just shifting by 1. Trunc-1, when rewritten, becomes $x = (1 + x)(1 + x^{-1})^{-1}$ which is satisfied and Trunc-2 becomes $1 + x(1 + y^{-1}) = 1 + (1 + x)(1 + y^{-1})$ which is not.)

The induced ordering, that is, the one obtained by defining $x \leq y$ iff $x \vee y = y$, is, of course, preserved under addition. And from that we may infer that it is reversed by negation: $x \leq y$ iff $x - (x + y) \leq y - (x + y)$ iff $-y \leq -x$. Hence negation must convert least upper bounds into greater lower bounds, yielding what can only be called "De Morgan's law": $-(x \wedge y) = (-x) \vee (-y)$. For a direct formula we have $x \wedge y = x - \lfloor x - y \rfloor$ (because $x \wedge y = -((-x) \vee (-y)) = -((-x) + \lfloor (-y) - (-x) \rfloor) = x - \lfloor x - y \rfloor$).

[117] That is, continuous functions $f : \mathbb{R}^n \to \mathbb{R}$ such that f agrees at each point with one of a finite set of (homogeneous) linear functions with integer coefficients.

proof that all LOAGs can be embedded in a product of TOAGs is—as is to be expected—essentially the same as it was for scales: it is necessary and sufficient to show that every subdirectly irreducible LOAG is a TOAG. Just as for scales (see Theorem 8.6 at p24), a consequence is that LOAGs are not just lattices but distributive lattices and that allows an easy proof that the lattices of congruences $\mathbf{Spec}(L)$, for a given LOAG L, is distributive and—as always for distributive lattices of congruences—therefore a spatial locale. When $\mathrm{Spec}(L)$ is viewed not as lattice but as a space, its points are the congruences for linear quotients [118] and if we specialize to $\mathrm{Max}(L)$, the congruences of simple quotients, the points are the congruences for Archimedean linear quotients.

Among the differences:

• *Not every non-trivial* LOAG *has a simple quotient.* Consider the TOAG of integral polynomials in one variable "ordered at infinity," that is, $f \leqq g$ iff $f(n) \leqq g(n)$ for almost all natural numbers n. For each quotient algebra there exists $d \in \mathbb{N}$ such that f and g name the same element in the quotient iff $\mathrm{degree}(f-g) < d$. None is maximal. There is no simple quotient. Another huge difference is that $\mathrm{Max}(-)$ is not a functor into the category of spaces; LOAG maps $K \to L$, when not epimorphic, do not induce functions $\mathrm{Max}(K) \to \mathrm{Max}(L)$, only partial functions.[119]

• $\mathrm{Max}(-)$ *is not a representable functor.* (Certainly every simple quotient may be embedded in the reals, but not uniquely: a non-trivial map $f : L \to \mathbb{R}$ certainly names a point in $\mathrm{Max}(L)$ but rf names the same point for every $r > 0$.) No "schizophrenic object."

• *For a* LOAG L *neither* $\mathrm{Spec}(L)$ *nor* $\mathrm{Max}(L)$ *need be compact.* It is the case that $\mathrm{Max}(L)$ is Hausdorff (as in Section 42, p95–97). The easiest non-compact examples to describe are the free LOAGs on infinitely many generators. (But $\mathrm{Max}(L)$ is compact for all finitely generated L.)

• *No non-trivial injectives.* Any LOAG may be embedded in another one in which it acquires an upper bound.[120] Hence no absolute retracts.

• $\mathrm{Max}(F_n)$, *where F_n is the free* LOAG *on n-generators, is not—as is the case for free scales—the n-cube. Far more interesting: it is the* $\boldsymbol{(n-1)}$***-sphere.*** It is not just the topology. To measure the distance between two maximal ideals use the probability that the two orderings they induce disagree whether an element of $\mathbb{Z}^n \subset F_n$ is positive. To be precise, for each $k \in \mathbb{N}$ define a k-walk to be a sequence of $k+1$ elements in $\mathbb{Z}^n$ such that each element agrees with the next on all but one coordinate, and that difference is 1. Let D_k be the proportion of all k-walks starting at the origin that end on an element on which the two orderings disagree. The probability of disagreement is $\lim_{k\to\infty} D_k$. Use de Moivre [121] and the rotational invariance of Gaussian distributions to establish that the result is the standard sphere geometry. (If we use Bernoulli [122] then with probability one we can compute the distance using the limiting frequency of disagreement on an endless random walk. Add Pólya to the mix and it doesn't even matter where we start.[123]) Those who insist on radians may multiply by π.

[118] Hence, as for scales, the points are the congruences of linear quotients and thus—again as for scales—every LOAG has a representation as the group of global sections of a sheaf of TOAGs.

[119] Since the domains will always be open we could obtain a functor by replacing it with *Scone*$(\mathrm{Max}(-))$ as described in Section 35 (p83–87).

[120] *e.g.*, given a LOAG L take the initial model of the equational theory in which each element of L appears as a constant plus one more constant b. Add to the theory of LOAGs the (variable-free) equations that describe structure of L and the fact that the b is their upper bound. If the map from L to the LOAG so constructed were not an embedding, that is, if there is a constant a such that there's a proof of $a = 0$ the join of the finite number of L-constants appearing in the proof could be substituted for b in each of the equations appearing in the proof. Hence a is already 0.

[121] Central limit theorem [122] Law of large numbers

[123] We could make this look more complicated by measuring the frequency that the two orderings disagree how a pair of random walks compare.

27. *Appendix:* Computational Complexity Issues

Finding which equations can be counterexampled in the theory of either scales or LOAGs is NP-complete. There is no substantive difference between the equational and universal Horn theory problems. Indeed, if we change scales to *linear* scales and LOAGs to TOAGs we obtain the result for the full first-order universal theory.

First, we observe that the "satisfaction" problem for Boolean algebras can be easily converted to a satisfaction problem of roughly the same size in scales. The proof would be straightforward if it were not the case that one of the variables in the formula $x \vee y = x \overset{\vee}{|} \widehat{\dot{x}\,|y}$ appears twice. An uncaring use of this formula would lead to an exponential growth in the length of the translation. Let $\mathcal{B}$ denote the Boolean algebra of extreme points in a scale. We avoid the problem by using:

27.1 LEMMA: *If* $x, y \in \mathcal{B}$ *then* $x \wedge y = x \overset{\wedge}{|} y$ *and* $x \vee y = x \overset{\vee}{|} y$.

(Check in any linear scale.)

Using these translations (plus the translation of negation into dotting) any Boolean term may be converted to a scale term of the same size. We are not done until we restrict the variables to $\mathcal{B}$. Given a scale term A on variables $v_1, v_2, \ldots, v_n$ any solution of the equation:

$$(v_1 \multimap \hat{v}_1)|(v_2 \multimap \hat{v}_2)|\cdots|(v_n \multimap \hat{v}_n)|A \quad = \quad \top$$

(associate at will) is a solution of $A = \top$ that necessarily lies entirely in $\mathcal{B}$.

The complementary problem is also convertible. Every equation in the theory of Boolean algebras is equivalent to an equation—of manageable size—in the theory of scales. (As in the satisfaction problem the proof would be straightforward if it were not the case that one of the variables in the formula for $x \vee y$ appears twice.)

For the conversion we will understand that $x \vee y$ is the term $x \overset{\vee}{|} \widehat{\dot{x}\,|y}$, the one in which y appears once and, dually, $x \wedge y$ is the term $x \overset{\wedge}{|} (\dot{x} \overset{\vee}{|} y)$, again the one in which y appears once. We will need a term $M\langle v_1, v_2, \ldots, v_n\rangle$ defined recursively by:

$$(n = 0) \;\Rightarrow\; (M = \top)$$

$$M\langle v_1, v_2, \ldots, v_n\rangle \quad = \quad (v_n \vee \dot{v}_n) \wedge M\langle v_1, v_2, \ldots, v_{n-1}\rangle$$

Given terms A and B in the signature of Boolean algebras we produce terms $\langle\!\langle A\rangle\!\rangle$ and $\langle\!\langle B\rangle\!\rangle$ such that $A = B$ is true for Boolean algebras iff $\langle\!\langle A\rangle\!\rangle = \langle\!\langle B\rangle\!\rangle$ is true for scales. The length of $\langle\!\langle A\rangle\!\rangle$ will be bound by a constant multiple of the length of A times the number of variables in A and B.

The conversion is defined recursively by the following rules in which $v_1, v_2, \ldots, v_n$ are the variables and M denotes $M\langle v_1, v_2, \ldots, v_n\rangle$.

$$\begin{aligned}
\langle\!\langle \top\rangle\!\rangle &= M \\
\langle\!\langle \bot\rangle\!\rangle &= \dot{M} \\
\langle\!\langle v_i\rangle\!\rangle &= \dot{M} \vee (v_i \wedge M) \\
\langle\!\langle \neg A\rangle\!\rangle &\doteq \langle\!\langle A\rangle\!\rangle^{\bullet} \\
\langle\!\langle A \vee B\rangle\!\rangle &= M \wedge [\dot{M} \lhd (\langle\!\langle A\rangle\!\rangle \mid \langle\!\langle B\rangle\!\rangle)] \\
\langle\!\langle A \wedge B\rangle\!\rangle &= \dot{M} \vee [M \lhd (\langle\!\langle A\rangle\!\rangle \mid \langle\!\langle B\rangle\!\rangle)]
\end{aligned}$$

Note that in a linear scale if any one of the variables is instantiated as $\odot$ then $M = \odot$, otherwise M and $\dot{M}$ are distinct. In either case, an inductive argument shows that $\langle\!\langle A\rangle\!\rangle$ is either M or $\dot{M}$ for every term A. Moreover $\langle\!\langle A \diamond B\rangle\!\rangle = \langle\!\langle A\rangle\!\rangle \diamond \langle\!\langle B\rangle\!\rangle$ where $\diamond$ is either lattice operation.

We may redo this construction for LOAGs instead of scales.

Define M by

$$\begin{aligned} (n = 0) &\Rightarrow (M = \top) \\ M\langle v_1, v_2, \ldots, v_n\rangle &= (\lfloor v_n\rfloor \vee \lfloor -v_n\rfloor) \vee M\langle v_1, v_2, \ldots, v_{n-1}\rangle \end{aligned}$$

and:

$$\begin{aligned} \langle\!\langle \top\rangle\!\rangle &= M \\ \langle\!\langle \bot\rangle\!\rangle &= -M \\ \langle\!\langle v_i\rangle\!\rangle &= (-M) \vee (v_i \wedge M) \\ \langle\!\langle \neg A\rangle\!\rangle &= -\langle\!\langle A\rangle\!\rangle \\ \langle\!\langle A \vee B\rangle\!\rangle &= M \wedge (\,\langle\!\langle A\rangle\!\rangle + \langle\!\langle B\rangle\!\rangle + M\,) \\ \langle\!\langle A \wedge B\rangle\!\rangle &= (-M) \vee (\,\langle\!\langle A\rangle\!\rangle + \langle\!\langle B\rangle\!\rangle - M\,) \end{aligned}$$

28. *Appendix:* Independence

The independence of all but the first two scale axioms is easy:

For the independence of the medial axiom, consider the set $\{-1, 0, +1\}$ with $x|y$ defined as "truncated addition," that is

$$x|y \;=\; -1 \vee (x + y) \wedge 1$$

We take $\dot{x} = -x$, $\hat{x} = x$ and $\top = 0$. All defining laws of scales hold except for the medial law $(+1|0) \mid (+1|-1) \neq (+1|+1) \mid (0|-1)$.

For the independence of the unital and constant laws consider the set $\{0, 1\}$ with ordinary multiplication for $x|y$ and the identity function for $\hat{x}$. If we take $\dot{x} = 1 - x$ and $\top = 1$ then every equation is satisfied except for $\bot\overset{\vee}{|}x = x$. If, instead, we take $\top = 0$ then every equation is satisfied except for $\top\overset{\wedge}{|}x = x$. If we take $\dot{x} = x$ and $\top = 1$ then every equation is satisfied except for the constancy of $\dot{x}\,|x$.

For the independence of the scale identity consider $\mathbf{I}\times\mathbf{I}$ with the standard product-algebra structure except for $\top$-zooming. The unital laws determine $\widehat{\langle x, y\rangle}$ only when x and y are both non-negative. We maintain all the laws except for the scale identity, therefore, if $\top$-zooming is standard on just that top quadrant.

As promised in [36] (p12) we can do better. The absorbing laws determine $\widehat{\langle x, y\rangle}$ only when x and y are both non-positive. We can maintain the minor-scale equations, therefore, by keeping the standard definition of $\widehat{\langle x, y\rangle}$ just on the top and bottom quadrants, (that is, the pairs $\langle x, y\rangle$ such that $xy \geqq 0$).

The first three uses of the scale identity were for $\overset{\vee}{\top} = \top$, $x = \overset{\vee}{x} \,|\, \hat{x}$ and $\widehat{x|\odot} = \hat{x}\,|\bot$ (the absorbing law, $\widehat{\bot|x} = \bot$, is a consequence of these). We may maintain the law of compensation by stipulating $\widehat{\langle x, y\rangle} = \langle x, y\rangle$ for $xy < 0$. Central distributivity requires a recursive definition. Given $\langle x, y\rangle$ such that $xy < 0$ let n be the largest integer such that there exist u, v with $\langle x, y\rangle = (\odot|)^n\langle u, v\rangle$. If $n = 0$ then define $\widehat{\langle x, y\rangle} = \langle x, y\rangle$. For $n > 0$ recursively define $\widehat{\langle x, y\rangle} = ((\odot|)^{n-1}\widehat{\langle u, v\rangle})|\langle\bot, \bot\rangle$. The scale identity itself fails (most easily seen by noting that $\top$-zooming no longer preserves order).

Also easy is the independence of the axioms for chromatic scales: if the support operation is constantly $\top$ then only the 1st equation, $\overline{\bot} = \bot$, fails; if it is the identity function then only the 2nd equation, $\hat{\overline{x}} = \overline{x}$, fails; if it is constantly $\bot$ then only the 3rd equation, $x \wedge \overline{x} = x$, fails; if $\overline{x} = \bot$ when $x = \bot$ else $\overline{x} = \top$ then the 4th equation, $\overline{x \wedge y} = \overline{x} \wedge \overline{y}$, fails when the scale is non-linear but only it fails.[124]

As promised, we can eliminate the 2nd equation by strengthening the 3rd equation to $x \wedge \dot{\overline{x}} = \bot$. Show first that $\dot{\overline{x}}$ is the Heyting negation, that is, $y \leqq \dot{\overline{x}}$ iff $y \wedge x = \bot$: if $y \leqq \dot{\overline{x}}$ then $y \wedge x \leqq \dot{\overline{x}} \wedge x = \bot$; if $y \wedge x = \bot$ then $y \vee \dot{\overline{x}} = (y \vee \dot{\overline{x}}) \wedge \top = (y \vee \dot{\overline{x}}) \wedge \dot{\overline{\bot}} = (y \vee \dot{\overline{x}}) \wedge \dot{\overline{y \wedge x}} = (y \vee \dot{\overline{x}}) \wedge (\dot{\overline{y}} \vee \dot{\overline{x}}) = (y \wedge \dot{\overline{y}}) \vee \dot{\overline{x}} = \bot \vee \dot{\overline{x}} = \dot{\overline{x}}$. This implies, in particular, that $\dot{\overline{x}}$ is an extreme point because $x \wedge \check{\dot{\overline{x}}} \leqq \check{x} \wedge \check{\dot{\overline{x}}} = (x \wedge \dot{\overline{x}})^\vee = \check{\bot} = \bot$ hence $\dot{\overline{x}} \leqq \check{\dot{\overline{x}}}$ and, consequently, $\dot{\overline{x}} \doteq \check{\dot{\overline{x}}}$. Since $\dot{\overline{x}}$ is an extreme point, so is $\overline{x}$. We obtain the original 3rd equation by $x \wedge \overline{x} = (x \wedge \overline{x}) \vee (\dot{\overline{x}} \wedge \overline{x}) = (x \wedge \dot{\overline{x}}) \vee \overline{x} = \bot \vee \overline{x} = \overline{x}$.

We have not yet established the independence of the idempotence and commutative laws. We already proved in [28] (p10) that the commutative law may be replaced with the single instance $\bot|\top = \top|\bot$ and we promised in [17] (p8) that we could remove the commutative law entirely by replacing the first unital law with $\dot{\bot} \overset{\wedge}{|}\, x = x$. To do so, first establish the left cancellation law using a different construction for dilatation, $(((\dot{a} \,|\, \dot{\bot})|x)^\wedge)^\vee$. Then $(((\dot{a} \,|\, \dot{\bot})\big|(a|x))^\wedge)^\vee = (((\dot{a} \,|\, a)\big|(\dot{\bot} \,|x))^\wedge)^\vee = (((\dot{\bot} \,|\bot)\big|(\dot{\bot} \,|x))^\wedge)^\vee = ((\dot{\bot}\big|(\bot|x))^\wedge)^\vee = (\bot|x)^\wedge = x$. Hence $a|x = a|y$ implies $x = y$ and just as in the derivation of full commutativity from the commutativity of $\top$ and $\bot$ we obtain dot-distributivity. Then obtain the involutory law from the second unital law written in full: $x = \bot \overset{\vee}{|}\, x = (((\bot|x)^{\bullet})^\wedge)^{\bullet} = ((\dot{\bot} \,|\, \dot{x})^\wedge)^{\bullet} = (\dot{x})^{\bullet}$. (Note in passing that we now have the original first unital law.) Finish as before by first showing: that $(\dot{x})^{\bullet} = x$ implies $x|\dot{x} = (\dot{x})^{\bullet}|\,\dot{x} = \odot$ hence the centrality of the center, $\odot|x = (x|\dot{x})\big|(x|x) = (x|x)\big|(\dot{x}|x) = x|\odot$. Finally use cancellation on $\odot\big|(x|y) = (\odot|x) \;\big|\; (\odot|y) = (\odot|x) \;\big|\; (y|\odot) = (\odot|y) \;\big|\; (x|\odot) = \odot\big|(y|x)$.

29. *Appendix:* Continuously vs Discretely Ordered Wedges

In Section 1 (p5–7) there appeared a quick and dirty procedure for computing the binary expansion of $f(x)$ where f is the unique interval–coalgebra-map from a given interval-coalgebra to the unit interval by iterating (forever):

$$\texttt{If } \check{x} = \top \texttt{ then emit } \text{“1”} \texttt{ and replace } x \texttt{ with } \hat{x}$$
$$\texttt{else emit } \text{“0”} \texttt{ and replace } x \texttt{ with } \check{x}.$$

A numerical analyst will object to the very beginning: how does one determine when an equality holds? There may be procedures that are guaranteed to detect when things are not equal (assuming, of course, that they are, indeed, not equal) but in analysis there tend not to be procedures that establish equality.[125]

[124] Similarly, the 1st, 3rd and 4th for support operations (see [95], p44) on rings are independent: if the support operation is constantly 1 then only $\overline{0} = 0$ fails; if it is constantly 0 then only $x\overline{x} = x$ fails; if $\overline{x} = 0$ then $x = 0$ else $\overline{x} = 1$ then $\overline{xy} = \overline{x}\,\overline{y}$ fails when the ring is not a domain but only it fails. (For the redundancy of the 2nd equation note first that it was not used to show that $x^2 = 0$ implies $x = 0$, hence $(1-\overline{x})x$ is 0 since its square is 0; finish with $(1-\overline{x})\overline{x} = (1-\overline{x})\overline{(1-\overline{x})}\overline{x} = (1-\overline{x})\overline{(1-\overline{x})x} = (1-\overline{x})\overline{0} = 0$.)

[125] For just one example, suppose the given interval-coalgebra is, itself, the unit interval but that we know an element x only by listening to its binary expansion. If that expansion happens to be .0 followed by all 1s we will never have enough information

Before considering computationally more realistic settings let us prove (in the classical setting) that the unit interval is the final interval coalgebra. Using binary expansions the interval coalgebra on $[0, 1]$ is described with an automaton with three states L, U, and initial state, M. It takes $\{0,1\}$-streams as input and produces $\{0,1\}$-streams as output:

Next State

	L	M	U
0	L	L	U
1	L	U	U

$\bot$-Zoom Output

	L	M	U
0	0		1
1	1		1

$\top$-Zoom Output

	L	M	U
0	0		0
1	0		1

The blanks in the output tables will be called **stammers**. The output streams will always be one digit behind the number of input digits.

We need a (dual) pair of definitions: define $\bot \ll \boldsymbol{x}$ if a finite iteration of $\bot$-zooming carries x to $\top$, and $\boldsymbol{x} \ll \top$ if a finite iteration of $\top$-zooming carries it to $\bot$. In $\mathbf{I}$ these are unneeded properties: $\ll$ coincides with $<$. [126] For any $\mathbf{I}$-valued coalgebra map, f, the two $\ll$-relations tell us how $f(x)$ is situated with respect to the center, $\odot$:

$$\begin{aligned} f(x) > \odot &\quad \text{iff} \quad \bot \ll \hat{x} \\ f(x) = \odot &\quad \text{iff} \quad \text{neither } \bot \ll \hat{x} \text{ nor } \overset{\vee}{x} \ll \top \\ f(x) < \odot &\quad \text{iff} \quad \overset{\vee}{x} \ll \top \end{aligned}$$

Note that $\bot \ll \hat{x}$ implies $\overset{\vee}{x} = \top$ (since $\hat{x} \neq \bot$) hence in the 1st case the procedure produces a stream of binary digits starting with 1 followed by the stream for $\hat{x}$ which is precisely what is demanded by $f(x) = (fx)^{\vee}|(fx)^{\wedge} = \top|f(\hat{x})$. The 3rd case is dual. In the 2nd case if $\overset{\vee}{x} = \top$ the procedure will produce a 1 followed by all 0s and if $\overset{\vee}{x} \neq \top$ a 0 followed by all 1s. That it produces one of these two streams (and it doesn't matter which) is just what is demanded by $f(x) = \odot$. [127]

For a computationally more realistic setting we are handed a guide to the needed modifications. The **Lawvere test** of a definition for the reals in a topos is that working in $\boldsymbol{Sh(X)}$, the category of sheaves on a space X, the definition yields the sheaf of continuous $\mathbb{R}$-valued functions on X, that is, the sheaf whose stalks are germs of continuous functions (as defined in the topos of sets) from X to $\mathbb{R}$. That experience leads us to view the sheaf of continuous $\mathbf{I}$-valued functions as the best candidate for the closed interval. We note immediately that the disjunctive coalgebra condition fails. Given continuous $g{:}\,X \to \mathbf{I}$ the "truth value" of the equation $\overset{\vee}{g} = \top$ (necessarily an open subset of X) is the interior of $g^{-1}[\odot, \top]$ and the value of $\hat{g} = \bot$ is the interior of $g^{-1}[\bot, \odot]$. Their union is not, in general, all of X. But it is a dense subset. Hence we replace the **discrete coalgebra condition**:

$$\bot = \hat{x} \quad \text{or} \quad \overset{\vee}{x} = \top$$

to conclude that $\overset{\vee}{x} = \top$. But this is not the most telling example (we know in this case that $\hat{x} = \bot$). Suppose that we know x only as the midpoint of a pair of binary expansions. If one sequence is the expansion of an arbitrary element y in the open unit interval and the other is the expansion of $1-y$ then we will never have enough information to know either $\overset{\vee}{x} = \top$ or $\hat{x} = \bot$.

[126] $<$ is the same as $\ll$ in a scale iff the scale is semi-simple.

[127] This argument works even in the intuitionistic setting if we hold on to the computationally unrealistic coalgebra condition $\overset{\vee}{x} = \top$ or $\hat{x} = \bot$ (as in the Cantor—rather than Dedekind—closed interval).

with the weaker **continuous coalgebra condition**:

$$\neg[\bot \ll x \text{ and } x \ll \top] \quad [128]$$

(We will understand that both the discrete and continuous coalgebra conditions entail that $\top$ and $\bot$ are zooming fixed-points.)

We must, however, capture the detectability of semiquations. Hence we replace the **apartness condition**:

$$\bot \neq \top$$

with this stronger **separation condition**:

$$\bot \neq x \quad \text{or} \quad x \neq \top \quad [129]$$

Keeping in mind that the truth values in $Sh(X)$ are open subsets of X, the extent to which any characteristic map, χ_A, is different from $\bot$ is the interior of A, the extent to which it is different from $\top$ is the "exterior" of A (the interior of its complement), hence $[\bot \neq \chi_A] \vee [\chi_A \neq \top]$ holds iff A is both open and closed.[130] The **continuously ordered wedge**, as opposed to the **discretely ordered wedge** [131] of X and Y is

$$\left\{ \langle x, y \rangle \; : \; \neg[x \ll \top \;\; \wedge \;\; \bot \ll y] \right\}$$

Its top is $\langle \top, \top \rangle$, its bottom $\langle \bot, \bot \rangle$. If either X or Y satisfies the separation condition, so does their ordered wedge. A continuous coalgebra structure on X is a map from X to the continuously ordered wedge of X with itself.[132]

---

[128] In footnote [9] (p6) about the modal operators $\Diamond$ and $\Box$ it was the continuous coalgebra condition we invoked when we wrote "No one allows simultaneously both $\Diamond\Phi \neq \top$ and $\Box\Phi \neq \bot$" (less than completely *possible/tenable/conceivable/allowed/foreseeable* but somewhat *necessary/certain/known/required/expected*). The discrete condition would have been the stronger "All insist upon $\Diamond\Phi = \top$ or $\Box\Phi = \bot$" (everything is either totally *possible/tenable/conceivable/allowed/foreseeable* or entirely *unnecessary/uncertain/unknown/unrequired/unexpected*.) The continuous coalgebra conditions sound realistic, the discrete do not. (The last pair—the "Bayesian modality pair"—is in lieu of a pair in which $\Box$ means *likely*—there seems to be no English word that works for the corresponding $\Diamond$, "not likely false." The pair should be viewed only as an approximation: among other ways of pronouncing this $\Box$ are *anticipated* and *foreseen*.)

[129] In the presence of De Morgan's law the two conditions are, of course, equivalent. In a topos the top and bottom of the subobject classifier Ω (or as Grothendieck called it, "the Lawvere object) are always apart; they are separate only when De Morgan holds throughout the topos. (Yes, they are apart—indeed separated—in the internal logic of the trivial topos.)

[130] Consider the sheaf of germs of all functions $f : X \to \mathbf{I}$, continuous or not, in the topos $Sh(X)$. We wish to find the largest separated subsheaf invariant under the zoom operators. So start by throwing away all germs that fail the separation condition. The trouble now is that the resulting sheaf is not closed under the zoom operators; so throw away all germs for which there is a zooming sequence α such that f^α fails the separation condition. The resulting sheaf is the sheaf of germs of continuous $\mathbf{I}$-valued functions. (Suppose f is not continuous at x. To find a discerning α start with the fact that there are values of f on arbitrarily small neighborhoods of x that are bounded away from fx. We may assume without loss of generality that those values are below fx. Let $\ell \in \mathbb{I}$ be such that $\ell < fx$ and for all neighborhoods of x there are values of f below ℓ. Let $u \in \mathbb{I}$ be such that $\ell < u < fx$. Let α be such that $\ell^\alpha = \bot$ and $u^\alpha = \top$. Then for any open $U \subseteq X$ such that $f^\alpha \neq \top$ on U it must be the case that $x \notin U$ and for any open $V \subseteq X$ such that $f^\alpha \neq \bot$ it must be the case, again, that $x \notin V$. That is $[\bot \neq f^\alpha] \vee [f^\alpha \neq \top]$ fails (because x is not in the union of $\{ y : \bot \neq f^\alpha y \}$ and $\{ y : f^\alpha y \neq \top \}$).)

[131] Sometimes "thick ordered wedge" as opposed to "thin ordered wedge."

[132] In footnote [2] (p4) we can replace the scale-algebras with interval coalgebras by adopting these intuitionistic modifications of the coalgebra definition to the classical setting. Given a set X with constants $\top$ and $\bot$ and unary operations whose values are denoted $\hat{x}$ and $\overset{\vee}{x}$ impose the conditions (where $\ll$ is as defined on p63):

$$\neg\left[\begin{array}{ccc} \bot \ll \hat{x} & \text{and} & \overset{\vee}{x} \ll \top \\ \bot \ll x & \text{or} & x \ll \top \end{array}\right]$$

The set, A, of sequences, $\mathbf{I}^{\mathbb{N}} = \prod_{\mathbb{N}} \mathbf{I}$, reduced by almost-everywhere equality does not satisfy either condition (consider, for example a sequence that is equal infinitely often to $\top$ and to $\bot$). But there is a largest subset that does, to wit, the set of sequences s such that for all zooming sequences α it is the case that

$$\neg\left[\begin{array}{ccc} \bot \ll s^{\alpha\wedge} & \text{and} & s^{\alpha\vee} \ll \top \\ \bot \ll s^\alpha & \text{or} & s^\alpha \ll \top \end{array}\right]$$

The resulting set is precisely the set of convergent sequences. If we now collapse to a point the set of its members such that $\neg(s \ll \top)$ and dually for $\neg(\bot \ll s)$ we obtain an interval coalgebra. Its unique coalgebra map to $\mathbf{I}$—as one must now have surely learned to expect—is *Lim*. And the same modifications work for defining limits of functions at a point in a space and for defining derivatives.

In preparation for establishing the nature of the final continuous coalgebra we'll say that an element c lies in the

$$\begin{array}{rl} \textbf{top open half} & \text{when } \bot \ll \hat{c} \\ \textbf{bottom open half} & \text{when } \overset{\vee}{c} \ll \top \end{array}$$

The top and bottom open halves do not form a cover (*e.g.*, $\odot \in \mathbf{I}$ is in neither of them) but the continuous coalgebra condition is precisely the condition that they are disjoint.[133] We do have a notion of "middle open half," to wit, when $\hat{c}$ is in the bottom open half and $\overset{\vee}{c}$ is in the top open half, that is, c lies in the

$$\textbf{middle open half} \quad \text{when } \overset{\vee}{\hat{c}} \ll \top \text{ and } \bot \ll \overset{\wedge}{\overset{\vee}{c}}$$

The critical lemma for the continuous case:

29.1 Critical Lemma: The Triumvirate of Open Halves
A continuously ordered wedge is the union of the three open halves: top, middle and bottom.

Because: we start with two instances of the separation condition

$$[\bot \ll \overset{\vee}{\hat{c}}] \vee [\overset{\vee}{\hat{c}} \ll \top] \quad \text{and} \quad [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \vee [\overset{\wedge}{\overset{\vee}{c}} \ll \top].$$

Clearly $\bot \ll \overset{\vee}{x}$ is equivalent with $\bot \ll x$ for any x and dually for $\hat{x} \ll \top$ and $x \ll \top$. We thus replace the two disjunctions with:

$$[\bot \ll \hat{c}] \vee [\overset{\vee}{\hat{c}} \ll \top] \quad \text{and} \quad [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \vee [\overset{\vee}{c} \ll \top].$$

The conjunction of these two disjunctions redistributes as a disjunction of four conjunctions:

$$[\bot \ll \hat{c}] \wedge [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \text{ or } [\bot \ll \hat{c}] \wedge [\overset{\vee}{c} \ll \top] \text{ or } [\overset{\vee}{\hat{c}} \ll \top] \wedge [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \text{ or } [\overset{\vee}{\hat{c}} \ll \top] \wedge [\overset{\vee}{c} \ll \top].$$

The second term is precisely what is prohibited by the definition of a continuously ordered wedge and we can weaken the first and last conjunctions [134] to obtain:

$$[\bot \ll \hat{c}] \quad \text{or} \quad [\overset{\vee}{\hat{c}} \ll \top] \wedge [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \quad \text{or} \quad [\overset{\vee}{c} \ll \top].$$

Exactly what we set out to prove.

[133] It's worth looking at the case when we're working in the topos of sheaves on a space X. The extent to which something is in a particular open half is given by an open subset of X.

[134] In fact they're not weakenings but equivalences: $[\bot \ll \hat{c}] \Rightarrow [\bot \ll \hat{c}] \wedge \left[[\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \vee [\overset{\wedge}{\overset{\vee}{c}} \ll \top] \right] \Rightarrow$
$\left[[\bot \ll \hat{c}] \wedge [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \right] \vee \left[[\bot \ll \hat{c}] \wedge [\overset{\wedge}{\overset{\vee}{c}} \ll \top] \right] \Rightarrow \left[[\bot \ll \hat{c}] \wedge [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \right] \vee \left[[\bot \ll \hat{c}] \wedge [\overset{\vee}{c} \ll \top] \right] \Rightarrow [\bot \ll \hat{c}] \wedge [\bot \ll \overset{\wedge}{\overset{\vee}{c}}].$

30. *Appendix:* Signed-Binary Expansions: the Contrapuntal Procedure and Dedekind Sutures

On July 31, 2000, I posted a note on the category net on how to obtain *signed* binary expansions for elements from such coalgebras. Five days later Peter Johnstone posted a note pointing out that—unlike the Dedekind-cut approach—my approach implicitly used the axiom of dependent choice.[135] In fact, I was using a very weak version of dependent choice [136] as will be explicated below. The approach can be modified without too much trouble (in a way easily seen to be equivalent to using Dedekind cuts) but that modification had not occurred to me at the time of the original publication. Because so many "infinite-precision" programmers are quite happy (wittingly or not) with dependent choice I did describe the approach in the printed version. We have much more inclusive motivation now that we have the necessary modification.

Every element of the standard interval has a representation of the form

$$\sum_{n=1}^{\infty} a_n 2^{-n}$$

where $a_n \in \{-1, 0, +1\}$. The sequence of a_ns is, of course, not unique: everything other than the two endpoints has infinitely many expansions—indeed, everything not a dyadic rational has continuously (or is it continuumly?) many.

The finite words on any alphabet may be viewed, of course, as a rooted tree and in the case at hand every vertex has three branches each with a label from $\{-1, 0, +1\}$.

It needn't be a tree. We get a diagram much easier to picture if we identify vertices when the paths that reach them are of the same length and name the same dyadic rational. Choose the "aspect ratio" to optimize the quality of the printed lines, that is, choose to make the oblique edges to be the nicest of Latex obliques, those with slope ± 1.

I like word-trees to grow to the right. (Why do so many like them to grow away from the light? Why do they like roots an top? Leaves on the bottom?) Combined with the choice of aspect we have a bonus: we don't need labels—slopes are all that anyone needs.

Then to make it fit, we handle the exponential growth of the number of edges by exponentially shrinking their length.

Presto! It becomes a wonderful illustration (appearing as our frontispiece) of just what the signed-binary sequences are doing. Because of its magical properties—most yet to be described—I have fallen into the habit of calling it the **Houdini diagram**.

If we view the diagram as a subset of the plane and take its closure, a boundary line is added to the right-hand of the tree. It's no less than the standard interval $\mathbf{I}$. If we follow a path from the far-left node we converge to the point in $\mathbf{I}$ named by the signed-binary expansion that uses the labels (the slopes) of the path. Indeed, as we travel that path the height above (or below) the central horizontal at each vertex is just what the signed-binary expansion describes at that point of path.

When writing signed-binary expansions we'll suppress the 1s and use the symbols

$$\text{“}-, \circ, +\text{”}$$

[135] Which axiom, it should be noted, is accepted by many constructive analysts.

[136] Which version is an easy consequence of the "disjunction property" that holds, for example, in the free topos with natural numbers object. See my Numerology in topoi. *Theory Appl. Categ.* 16 (2006) *No. 19,* p522–528

Given an object C with separated elements $\bot$ and $\top$ and self-maps whose values are denoted $\overset{\vee}{x}, \hat{x}$ satisfying the continuous coalgebra condition we seek a procedure that delivers for each $c \in C$ a signed-binary expansion.

In the last Section, Theorem 29.1 said that a continuously ordered wedge is the union of the three open halves: bottom, middle and top.

When c is in the top open half we want to emit "+" and replace c with $\hat{c}$, when in the bottom open half we want to emit "−" and replace c with $\overset{\vee}{c}$. And when c is in the middle open half we want to emit "o" and replace c with its its mid-zoom $\overset{\leftrightarrow}{c}$.

Whoops.

We don't have a mid-zoom operation.

But the continuously ordered wedge $C \vee C$ does have a mid-zoom and that solves the problem. For a moment let T be an arbitrary endofunctor on a category and $F \to TF$ a final T-coalgebra. Given an arbitrary coalgebra $g : C \to TC$ we have another coalgebra $Tg : TC \to T^2C$. (Note that it's an absolute tautology that $g : C \to TC$ is a map of coalgebras.) If we can describe a coalgebra map $TC \to F$ then $C \to TC \to F$ is the unique coalgebra map from C to F. [137] In the case at hand the induced bottom-zoom function $TC \to TC$ sends $\langle u, v\rangle$ to $\langle u \overset{\vee}{,} v\rangle = \langle \overset{\vee}{u}, \hat{u}\rangle$ and the induced top-zoom function sends $\langle u, v\rangle$ to $\langle u \overset{\wedge}{,} v\rangle = \langle \overset{\vee}{v}, \hat{v}\rangle$. We define the **mid-zoom** to be the function that sends $\langle u, v\rangle$ to $\boldsymbol{\langle u \overset{\leftrightarrow}{,} v\rangle = \langle \hat{u}, \overset{\vee}{v}\rangle}$. [138]

The critical prcperty of mid-zooming is: $\langle u \overset{\wedge}{\overset{\leftrightarrow}{,}} v\rangle = \langle u \overset{\vee}{\overset{\wedge}{,}} v\rangle$ and $\langle u \overset{\vee}{\overset{\leftrightarrow}{,}} v\rangle = \langle u \overset{\wedge}{\overset{\vee}{,}} v\rangle$. [139]

With the mid-zoom in hand we start again:

The triumvirate says that for any c in an interval coalgebra C

$$[\bot \ll \hat{c}] \text{ or } [\overset{\vee}{\hat{c}} \ll \top] \wedge [\bot \ll \overset{\wedge}{\overset{\vee}{c}}] \text{ or } [\overset{\vee}{c} \ll \top].$$

Hence for any $\langle u, v\rangle \in C \vee C$:

$$[\langle \bot, \bot\rangle \ll \langle u \overset{\wedge}{,} v\rangle] \text{ or } [\langle u \overset{\vee}{\overset{\wedge}{,}} v\rangle \ll \langle \top, \top\rangle] \wedge [\langle \bot, \bot\rangle \ll \langle u \overset{\wedge}{\overset{\vee}{,}} v\rangle] \text{ or } [\langle u \overset{\vee}{,} v\rangle \ll \langle \top, \top\rangle].$$

which translates to:

$$[\langle \bot, \bot\rangle \ll \langle \overset{\vee}{v}, \hat{v}\rangle] \text{ or } [\langle \overset{\vee}{\overset{\vee}{v}}, \overset{\wedge}{\overset{\vee}{v}}\rangle \ll \langle \top, \top\rangle] \wedge [\langle \bot, \bot\rangle \ll \langle \overset{\vee}{\hat{u}}, \overset{\wedge}{\hat{u}}\rangle] \text{ or } [\langle \overset{\vee}{u}, \hat{u}\rangle \ll \langle \top, \top\rangle].$$

Note that the nth iteration of $\bot$-zooming turns x into $\top$ iff it turns $\langle \overset{\vee}{x}, y\rangle$ into $\langle \top, \top\rangle$ for any y. Hence $\langle \bot, \bot\rangle \ll \langle \overset{\vee}{v}, \hat{v}\rangle$ iff $\bot \ll v$ and dually, $\langle \overset{\vee}{u}, \hat{u}\rangle \ll \langle \top, \top\rangle$ iff $u \ll \top$. When we specialize x to $\hat{u}$ we obtain $\langle \bot, \bot\rangle \ll \langle \overset{\vee}{\hat{u}}, \overset{\wedge}{\hat{u}}\rangle$ iff $\bot \ll \hat{u}$ and dually $\langle \overset{\vee}{\overset{\vee}{v}}, \overset{\wedge}{\overset{\vee}{v}}\rangle \ll \langle \top, \top\rangle$ iff $\overset{\vee}{v} \ll \top$. All of which says that if $\langle u, v\rangle \in C \vee C$ then the triumvirate of the open halves is equivalent to:

$$[\bot \ll v] \text{ or } [\hat{v} \ll \top] \wedge [\bot \ll \overset{\vee}{u}] \text{ or } [u \ll \top].$$

[137] Has this fact ever been used before?

[138] The verification that the values of the mid-zoom all lie in the continuously ordered wedge uses again that the zoom functions fix the endpoints. Indeed, for any pair $g, h : C \to C$ that fix $\bot$ and $\top$ it is the case that if $\langle u, v\rangle$ is in the continuously ordered wedge then so is $\langle g(u), h(v)\rangle$.

[139] Besides the evidence from the Houdini diagram we have a proof: $\langle u \overset{\wedge}{\overset{\leftrightarrow}{,}} v\rangle = \overset{\wedge}{\langle \hat{u}, \overset{\vee}{v}\rangle} = \langle \overset{\vee}{\overset{\vee}{v}}, \overset{\wedge}{\overset{\vee}{v}}\rangle = \overset{\vee}{\langle \overset{\vee}{v}, \hat{v}\rangle} = \langle u \overset{\vee}{\overset{\wedge}{,}} v\rangle$.

Given an element c in a continuously ordered wedge let $\langle u,v\rangle = \langle \overset{\vee}{c}, \hat{c}\rangle$ and iterate (forever) the non-deterministic parallel **contrapuntal procedure**:

$$\left[\begin{array}{c}\texttt{If } \bot \ll v\\ \texttt{emit } \text{“+”};\\ \texttt{replace } \langle u,v\rangle \texttt{ with}\\ \langle u\overset{\wedge}{,}v\rangle = \langle \overset{\vee}{v}, \hat{v}\rangle.\end{array}\right] \Big\| \left[\begin{array}{c}\texttt{If } \bot \ll \hat{u} \texttt{ and } \overset{\vee}{v} \ll \top\\ \texttt{emit } \text{“o”};\\ \texttt{replace } \langle u,v\rangle \texttt{ with}\\ \langle u\overset{\leftrightarrow}{,}v\rangle = \langle \hat{u}, \overset{\vee}{v}\rangle.\end{array}\right] \Big\| \left[\begin{array}{c}\texttt{If } u \ll \top\\ \texttt{emit } \text{“−”};\\ \texttt{replace } \langle u,v\rangle \texttt{ with}\\ \langle u\overset{\vee}{,}v\rangle = \langle \overset{\vee}{u}, \hat{u}\rangle.\end{array}\right]$$

The Houdini diagram brought one important feature of the contrapuntal procedure to my attention (13 years late). We'll say that a signed-binary stream is an **OK stream** if it it does not have a **bad tail**, to wit, an infinite stream of all +s or all −s but is not all +s or all −s (that is, it is not one of the two **edge streams**, to wit, the unique streams for $\top$ and $\bot$).

30.2 LEMMA: *Streams produced by the contrapuntal procedure are OK streams.*

(This wonderful fact turns out to solve a number of irritating problems.) Suppose, first, that an output stream were to end with the bad tail o+++···. That first o says that the procedure determined that it was at a node where the second alternative is viable, that is, where the rest of the output steam lies in the middle open half. Any initial segment of that infinite stream does just that. But the bad tail in question converges to a point not in the middle *open* half, hence that first o would not have been possible. A similar argument holds for the possibility of +−−− ···. That first + says that the first alternative is viable, that is, the element being described lies in the top open half. But the bad tail in question describes an element not in the top open half.[140]

We can use even more. First, though, a definition. Note that in the Houdini diagram a node can have 0,1 or 2 incoming edges. We need a name, in particular for the case of a single incoming edge. We'll call them **monodes**. A monode is named either by a finite word ending in o or it is a node on one of the two outer edges (the "one" excludes the root).

It turns out to be important that not only can we eliminate bad tails but we can eliminate all streams with only finitely many monodes. We'll call those with infinitely many monodes **good streams**.

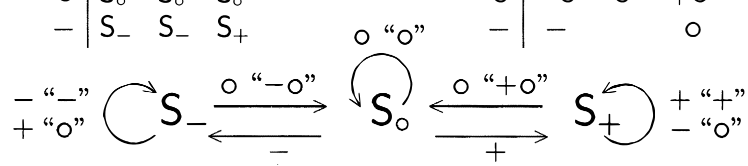
The following finite automaton converts any stream without a bad tail into a good stream that names the same element of $\mathtt{I}$. The states are $\mathsf{S}_-, \mathsf{S}_\circ$ (the initial state) and S_+.

Next State

	S_-	$\mathsf{S}_\circ$	S_+
+	S_-	S_+	S_+
o	$\mathsf{S}_\circ$	$\mathsf{S}_\circ$	$\mathsf{S}_\circ$
−	S_-	S_-	S_+

Output

	S_-	$\mathsf{S}_\circ$	S_+
+	o		+
o	−o	o	+o
−	−		o



[140] It is worth checking that even if we were to start with a stream procedure the output will not end with a bad tail. Suppose the input sequence is $c = -++++\cdots$. Then $\langle \overset{\vee}{c}, \hat{c}\rangle = \langle +++\cdots, ----\cdots\rangle$ and only the second alternative holds hence the output is o and—since both $\overset{\vee}{c}$ and $\hat{c}$ are zooming fixed-points—all further further outputs will be o. Suppose—instead—that the input sequence is $c = \text{o}++++\cdots$. Then $\langle \overset{\vee}{c}, \hat{c}\rangle = \langle +++\cdots, \text{ooo}\cdots\rangle$ and only the first alternative holds hence the output is + and the pair is replaced, again, with $\langle +++\cdots, ----\cdots\rangle$ and all further outputs will, again, be o.
See Section 43 (p97–107) for a way of making a much simpler finite automaton that removes bad tails.

The states may be interpreted as follows: in state $\mathsf{S}_\circ$ the the input stream (so far) is numerically equal to the present output stream. Whenever we leave $\mathsf{S}_\circ$ a stammer occurs and the machine moves to either S_+ or S_-; in S_+ the input stream is larger than the output; in S_- it's smaller. Whenever we return to $\mathsf{S}_\circ$ a **stutter** occurs: two output digits. The machine is never more than one output digit behind the number of input pairs (that is, between every pair of stammers there's a stutter).[141]

Let $\alpha \in \{-, \circ, +\}^*$ be a finite word of signed binary digits. The **scope** of the node defined by α, $\boldsymbol{Sc(\alpha)}$, is the regular open subinterval of $\mathbb{I}$ with endpoints $\alpha{-}{-}{-}{-}\cdots$ and $\alpha{+}{+}{+}{+}\cdots$. It is none other than the subset of $\mathbb{I}$ reachable by streams without bad tails that are continuations of α. (Note that any *regular* open subinterval of $\mathbb{I}$ that shares an endpoint with $\mathbb{I}$ must include that endpoint.) This is the critical use of the absence of bad tails. If all streams are allowed we reach a closed subinterval from a given node, hence I saw no good way of using the contrapuntal procedure—the very foundation of the finality of the standard closed interval—as a tool for the construction of the standard interval.

The set of nodes, $\mathcal{N}$, is isomorphic to the set of natural numbers.[142] Given an interval coalgebra C and an element $c \in C$ we obtain the subset of nodes reachable by the contrapuntal procedure with c as input. We wish to characterize those subsets.

In a topos with natural numbers object we can view the output, therefore, as a map $D : \mathcal{N} \to \Omega$. Given a finite word α of signed binary digits we may view $D(\alpha)$ as the "truth-value" that the node named by α is in the subset (described by D). We'll call $\boldsymbol{D(\alpha)}$ the **domain** of the node named by D, which function will be called the **domain function**.

The Lawvere test says that we need a condition on domain functions so that in $Sh(X)$ domain functions are in one-to-one correspondence with continuous maps $f : X \to \mathbb{I}$.

The only known way of eliminating the evil of bad tails is to use the plenitude of monodes.

We'll use the conventions that λ names the empty word and that $\boldsymbol{\alpha \prec \beta}$ means that for $\beta = \alpha\gamma$ for some non-empty γ that ends with a monode:

The output of the contrapuntal procedure can be described as a map $D : \mathcal{N} \to \Omega$ satisfying the two conditions:

$$\begin{aligned} D(\lambda) &= \mathsf{True} \\ D(\alpha) \wedge D(\beta) &= \bigvee \{ D(\delta) : \alpha \prec \delta \wedge \beta \prec \delta \} \end{aligned}$$

(Keep in mind: $D(\alpha) = D(\alpha) \wedge D(\alpha)$.) We'll call such functions **Dedekind sutures**.

---

[141] A non-stuttering machine (with 4 states, one of which, I, is strict initial) is available:

Next State

	I	S_-	$\mathsf{S}_\circ$	S_+
$+$	S_+	S_-	S_+	S_+
$\circ$	$\mathsf{S}_\circ$	$\mathsf{S}_\circ$	$\mathsf{S}_\circ$	$\mathsf{S}_\circ$
$-$	S_-	S_-	S_-	S_+

Output

	I	S_-	$\mathsf{S}_\circ$	S_+
$+$		$\circ$	$\circ$	$+$
$\circ$		$-$	$\circ$	$+$
$-$		$-$	$\circ$	$\circ$

A stammer occurs at the very beginning, thereafter it is always exactly one output digit behind the number of input pairs; alternatively, a machine that doesn't stammer until it has to:

Next State

	I	S_-	$\mathsf{S}_\circ$	S_+
$+$	S_+	S_-	S_+	S_+
$\circ$	I	$\mathsf{S}_\circ$	$\mathsf{S}_\circ$	$\mathsf{S}_\circ$
$-$	S_-	S_-	S_-	S_+

Output

	I	S_-	$\mathsf{S}_\circ$	S_+
$+$		$\circ$	$\circ$	$+$
$\circ$	$\circ$	$-$	$\circ$	$+$
$-$		$-$	$\circ$	$\circ$

[142] One choice of canonical names is the Kleene-regular set of words $\{-,+\}^*\{\circ\}^*$, that is, finite $\circ$-free words followed by finite strings of $\circ$s. (Every dyadic rational strictly between -1 and $+1$ is described by a unique finite $\circ$-free word.) For a specific isomorphism define $\mathbb{N} \to \mathcal{N}$ by sending $(2n+1)2^m - 1$ to the node reached by the word $(fn)\circ^m$ where $f : \mathbb{N} \to \{+,-\}^*$ is the unique function such that $f0$ is the empty word, $f(2n+1) = (fn)+$ and $f(2n+2) = (fn)-$. Cf. [145] (p71).

In $Sh(X)$ we can rewrite the conditions (Ω can be taken as the set of open subsets of X):

$$\begin{aligned} D(\lambda) &= X \\ D(\alpha) \cap D(\beta) &= \bigcup \{ D(\delta) \,:\, \alpha \prec \delta \wedge \beta \prec \delta \} \end{aligned}$$

Given continuous $f : X \to [-1,+1]$ define $D : \mathcal{N} \to \Omega$ by $D(\alpha) = f^{-1}(Sc(\alpha))$. The verification that such is a Dedekind suture is routine.

Given a Dedekind suture $D : \mathcal{N} \to \Omega$ define $f : X \to [-1,+1]$ by taking $f(x)$ to be the unique element in $f(x)$ to be the unique element in $\bigcap \{ Sc(\alpha) \,:\, x \in D(\alpha) \}$.

The fact that we do obtain an element, that is, the fact that the intersection of the good scopes is non-empty, that fact is a critical use of the plenitude of monodes. Suppose their intersection were empty. Replace each interval with its closure; the unique element in their intersection is necessarily one of the new endpoints, hence necessarily a dyadic rational and one not equal to $\bot$ or $\top$; but the only good streams converging to such are eventually all zero and the dyadic rational is not a new endpoint; indeed it's eventually the center of each of the scopes.

For the continuity of f we use the easily verified facts that 1) the good-stream scopes form a basis for the topology of the standard interval and that 2) $f^{-1}(Sc(\alpha)) = D(\alpha)$ for all α.

If one starts with a continuous function, f, and proceeds to the construction of its domain function D, it is easy to verify that the inverse images of the good-stream scopes of the function constructed from D are, of course, the same as the inverse images of f. And in the other direction, it is also routine that if one starts with a Dedekind suture, D, and constructs a continuous function $f : X \to \mathrm{I}$ then the domain function of f is D.[143]

Since we described the automata for zooming automata in the discrete case, we close this section with the automata for the signed-binary-digit setting. The diagram below is for $\top$-zooming; negate the 6 signatures for $\bot$-zooming.

Next State

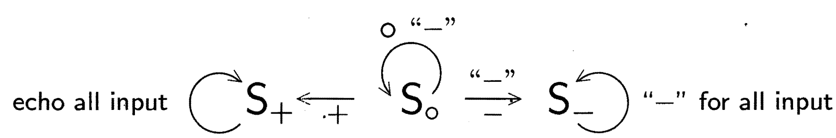
	S_-	$S_\circ$	S_+
$+$	S_-	S_+	S_+
$\circ$	S_-	$S_\circ$	S_+
$-$	S_-	S_-	S_+

$\top$-Zoom Output

	S_-	$S_\circ$	S_+
$+$	$-$		$+$
$\circ$	$-$	$-$	$\circ$
$-$	$-$	$-$	$-$

$\bot$-Zoom Output

	S_-	$S_\circ$	S_+
$+$	$+$	$+$	$+$
$\circ$	$\circ$	$+$	$+$
$-$	$-$		$+$



In these machines at most one stammer occurs; restated, the output is never more than one digit behind the input.[144]

It's worth noting that no automaton, finite or not, can compute midpoints (or, as usually pointed out, sums) in the context of streams of standard (unsigned) binary digits. If at any point only heterogeneous pairs of standard digits (one 0, one 1) have been heard then we do

[143] The "disjunction property" obviates the need for dependent choice. Given a space X and a continuous coalgebra C in the category of sheaves $Sh(X)$ and a partial section $U \to C$ we seek continuous $f : U \to \mathrm{I}$. (Dependent choice holds only when X is totally disconnected.) For each $x \in X$ we are prompted to pass to the category of "micro-sheaves" at x, to wit, the result of identifying things when they agree when restricted to some neighborhood of x. In that category we are entitled to view the partial section as global, a map from 1 to C. What we gain is the "disjunction property": if the disjunction of two (or, more to the point, three) sentences is true then one of them is already true. We may now repeat the above procedure to obtain an unending stream of signed binary digits. Continuity is left as an exercise.

[144] There are four-state automata with strict initial states, I, that always stammer at the first input digit and never thereafter:

not know whether the result will be in the upper or lower open half of the unit interval and if we are restricted to the digits 0 and 1 we can not specify its first digit. If, perchance, all digit-pairs are heterogeneous we will never be able to compute the first digit.

There is a remarkably simple automaton, on the other hand, for midpointing signed binary digits. See Section 43, p97–107 for the scale-structure automata of signed binary expansions.

31. *Appendix:* Dedekind Cuts

If we wish to avoid the axiom of dependent choice, one approach is to use Dedekind cuts. As previously noted, if experience with topoi is any guide we know in advance what $\mathtt{I}$ should turn out to be in the category of sheaves over a topological space X, to wit, the sheaf of continuous $\mathtt{I}$-valued functions on X, that is, the sheaf whose stalks are germs of continuous functions (as defined in the topos of sets) from X to $\mathtt{I}$. There are a number of ways of define a Dedekind cut—all of them equivalent in the category of sets—but not all give give the right answer in a category of sheaves.

When working with $\mathbb{I}$ rather than $\mathbb{D}$ it is convenient to define Dedekind cuts as subsets not of $\mathbb{I}$ but of its "interior," $\mathring{\mathbb{I}}$, the result of removing the two endpoints.[145]

We say that $L \subseteq \mathring{\mathbb{I}}$ is a **downdeal** if:

$$\ell \in L \;\Rightarrow\; \forall_{\ell' < \ell}\, \ell' \in L$$

and it is an **open downdeal** if, further:

$$\ell \in L \;\Rightarrow\; \exists_{\ell' > \ell}\, \ell' \in L$$

We say that U is an **updeal** if

$$u \in U \;\Rightarrow\; \forall_{u' > u}\, u' \in U$$

and it is an **open updeal** if, further

$$u \in U \;\Rightarrow\; \exists_{u' < u}\, u' \in U$$

In the case of sheaves on X we may reinterpret L and U as a families of open subsets of X indexed by elements of $\mathring{\mathbb{I}}$ (that is, $L_{\ll}$ gives the "extent to which $\ell \in L$"). The conditions then rewrite to:

$$L_\ell \;=\; \bigcup_{\ell' > \ell} L_{\ell'} \qquad\qquad U_u \;=\; \bigcup_{u' < u} U_{u'}$$

Next State

	I	L	M	U
+	U	L	U	U
o	M	L	M	U
−	L	L	L	U

$\top$-Zoom Output

	I	L	M	U
+		+	+	+
o		o	+	+
−		−	+	+

$\bot$-Zoom Output

	I	L	M	U
+		−	−	+
o		−	−	o
−		−	−	−

Six states are required for a mid-zoom machine and its unique stammer is also at the beginning:

Next State

	I	L_2	L_1	M	U_1	U_2
+	U_1	L_2	M	M	U_2	U_2
o	M	L_2	L_1	M	U_1	U_2
−	U_1	L_2	L_2	M	M	U_2

Output

	I	L_2	L_1	M	U_1	U_2
+		−	−	+	+	+
o		−	−	o	+	+
−		−	−	−	+	+

[145] Bear in mind that $\mathbb{I}$ and $\mathring{\mathbb{I}}$ are isomorphic—as objects—to the natural numbers $\mathbb{N}$. For a specific isomorphism define $h : \mathbb{N} \to \mathring{\mathbb{I}}$ to be the unique function such that $h(0) = 0$, $h(2n+1) = \frac{h(n)+1}{2}$ and $h(2n+2) = \frac{h(n)-1}{2}$. Cf. [142] (p69).

Note that for any downdeal L, open or not, the largest open downdeal contained therein is

$$\mathring{L} = \left\{ \ell : \exists_{\ell' \in L} \, \ell' > \ell \right\}$$

and the largest open updeal contained in an updeal U is a

$$\mathring{U} = \left\{ u : \exists_{u' \in U} \, u' < u' \right\}$$

When working in sheaves these correspond to

$$\mathring{L}_\ell = \bigcup_{\ell' > \ell} L_{\ell'} \qquad \mathring{U}_u = \bigcup_{u' < u} U_{u'}$$

For any lower semicontinuous map $f : X \to \mathbf{I}$ we obtain an open downdeal by defining $L_\ell = f^{-1}(\ell, \top] \cap \mathring{\mathbb{I}}$ and for any upper semicontinuous g we obtain an open updeal with $U_u = g^{-1}[\bot, u) \cap \mathring{\mathbb{I}}$. Conversely, given an open downdeal, L, define $f(x)$ to be $\sup \{ \ell : x \in L_\ell \}$ and given an open updeal, U, define $g(x) = \inf \{ u : x \in U_u \}$. It is easy to check that these assignments establish a correspondence between upper/lower semicontinuous functions and open up/down-deals.

In the classical setting there are several ways of defining $\mathbf{I}$: as the set of open downdeals; as the set of open updeals; as the set of pairs $\langle L, U \rangle$ where L is maximal among those downdeals disjoint from U and U is maximal among those disjoint from L. In the more general setting, none of these are guaranteed to satisfy the separation condition: on any space X and open set V take L_ℓ to be constantly equal to V. The extent to which L is $\top$ is V and the extent to which it is different from $\top$ is $\neg V$, to wit, its exterior (defined as the largest open set disjoint from V, the interior of its complement). The extent to which L is not $\bot$ is $\neg\neg V$ (the interior of the closure of V). The union $\neg V \cup \neg\neg V$ fails, in general, to be all of X. Note that the maximal open updeal disjoint from L is U where U_u is constantly $\neg V$. The maximal open downdeal disjoint from U is $\neg V$. If we take any "adjoint pair" (p56) of semicontinuous maps $\langle f, g \rangle$ and $x \in X$ such that $f(x) < g(x)$ we may let $r = f(x)|g(x)$ and apply a suitable power of the dilatation at r to obtain an adjoint pair $\langle (r\triangleleft)^n f, (r\triangleleft)^n g \rangle$ that will recreate the same sort of pathology. We need, in other words. a condition on $\langle L, U \rangle$ that will force $f = g$.

Given any function $h : X \to \mathbf{I}$ there is a maximal lower semicontinuous $f : X \to \mathbf{I}$ below h and a minimal upper semicontinuous g above h. (To obtain f define a downdeal L by first taking L_ℓ to be the interior of $h^{-1}(\ell, \top]$ and then replacing L_ℓ with $\mathring{L}_\ell$.) Of course a function both upper and lower semicontinuous is plain continuous.

A **Dedekind cut** on $\mathring{\mathbb{I}}$ is a pair of subsets $\langle L, U \rangle$ such that:

L is an open downdeal:	$\ell \in L$	$\Rightarrow$	$[\forall_{\ell' < \ell} \, \ell' \in L] \; \& \; [\exists_{\ell' > \ell} \, \ell' \in L]$
U is an open updeal:	$u \in U$	$\Rightarrow$	$[\forall_{u' > u} \, u' \in U] \; \& \; [\exists_{u' < u} \, u' \in U]$
L and U disjoint:	$q \in \mathring{\mathbb{I}}$	$\Rightarrow$	$\neg[(q \in L) \; \& \; (q \in U)]$
L and U almost cover:	$\ell < u$	$\Rightarrow$	$[\ell \in L]$ or $[u \in U]$

(L and U each determines the other: given L then $U = \{ u : \exists_{u' < u} u' \notin L \}$. The conditions are, in fact, redundant: the 3rd and 4th conditions imply that L is a downdeal and U an updeal (but not the openness condition).)

In the case of sheaves on X the conditions rewrite to:

$$\begin{aligned} U_u &= \bigcup_{u'<u} U_{u'} \\ L_\ell &= \bigcup_{\ell'>\ell} L_{\ell'} \\ L_q \cap U_q &= \emptyset \\ \ell < u &\Rightarrow L_\ell \cup U_u = X \end{aligned}$$

For any continuous $f{:}X \to \mathbf{I}$ we obtain such a cut be defining $L_\ell = f^{-1}(\ell, \top]$ and $U_u = f^{-1}[\bot, u)$. All Dedekind cuts so arise: given the L_ℓs and U_us define $f: X \to \mathbf{I}$ by $f(x) = \inf\{\, u : x \in U_u \,\}$ and verify that f is continuous (the key observation is that the closure of U_u is contained in $U_{u'}$ whenever $u < u'$, hence $\{\, x : f(x) \leqslant u \,\} = \bigcap_{u'>u} U_{u'} = \bigcap_{u'>u} \overline{U}_{u'}$ is closed and, dually, $\{\, x : f(x) \geqslant u \,\}$ is open).

We define $\mathbf{I}$ to be the set of such Dedekind cuts. The bottom cut is $\langle \emptyset, \mathring{\mathbb{I}} \rangle$; the top cut is $\langle \mathring{\mathbb{I}}, \emptyset \rangle$.

Note that $\langle L, U\rangle \neq \bot$ iff L is non-empty and, dually, $\langle L, U\rangle \neq \top$ iff U is non-empty. The almost-cover condition for the case $\bot < \top$ is precisely the separation condition for the set of Dedekind cuts.

Define

$$\begin{aligned} \langle L \overset{\vee}{,} U\rangle &= \langle\, \{\overset{\vee}{\ell} : \odot > \ell \in L\},\ \{\overset{\vee}{u} : \odot > u \in U\rangle \,\} \\ \langle L \overset{\wedge}{,} U\rangle &= \langle\, \{\overset{\wedge}{\ell} : \odot < \ell \in L\},\ \{\hat{u} : \odot < u \in U\rangle \,\} \end{aligned}$$

The Dedekind-cut conditions for $\langle L \overset{\vee}{,} U\rangle$ and $\langle L \overset{\wedge}{,} U\rangle$ are pretty routine. (For the almost-cover condition, given $\ell < u$ use the scale structure on $\mathbb{I}$: since $(\top|\ell) < (\top|u)$ we know that either $(\top|\ell) \in L$ or $(\top|u) \in U$. In the first case $\odot \leqslant (\top|\ell)^\wedge \in L$ and in the second case $\odot < (\top|u)^\wedge \in U$.)

The disjointness condition for the case $\odot$ is precisely the continuous coalgebra condition for these operations.

Given an interval coalgebra, C, satisfying the separation and continuous coalgebra condition, and given $c \in C$, we wish to construct a Dedekind cut $\langle L(c), U(c)\rangle$. By a **zooming sequence** is meant an element of the free monoid on two generators, $\top$-zooming and $\bot$-zooming.

$\ell \in L(c)$ *iff there is a zooming sequence* α *such that* $\ell^\alpha = \bot$ *and* $\neg\neg[c^\alpha = \top]$.

$u \in U(c)$ *iff there is a zooming sequence* α *such that* $u^\alpha = \top$ *and* $\neg\neg[c^\alpha = \bot]$.

We need to verify the conditions for a Dedekind cut.

Before doing so let us pause to collect a few easily verified observations in the intuitionistic setting. For any function, f, it is, of course, trivial that $(x = y) \Rightarrow (fx = fy)$. Because negation is contravariant we also have $(fx \neq fy) \Rightarrow (x \neq y)$ and $\neg\neg(x = y) \Rightarrow \neg\neg(fx = fy)$. We will apply these trivial observations to the case when f is a zooming sequence and incorporate the fact that $\top$ and $\bot$ are fixed-points. Hence

$$\begin{aligned} x = \top &\Rightarrow x^\alpha = \top \\ x = \bot &\Rightarrow x^\alpha = \bot \\ x^\alpha \neq \top &\Rightarrow x \neq \top \\ x^\alpha \neq \bot &\Rightarrow x \neq \bot \\ \neg\neg[x = \bot] &\Rightarrow \neg\neg[x^\alpha = \bot] \\ \neg\neg[x = \bot] &\Rightarrow \neg\neg[x^\alpha = \bot] \end{aligned}$$

We will also use these trivial consequences of the apartness of $\top$ and $\bot$:

$$\begin{array}{lll} \neg\neg[x=\bot] & \Rightarrow & x \neq \top \\ \neg\neg[x=\top] & \Rightarrow & x \neq \bot \end{array} \quad [146]$$

And we will freely use all sorts of nice properties enjoyed by $\mathbb{I}$ (including the discrete coalgebra condition).

Not so trivial is this critical lemma:

31.1 LEMMA: *For $c \in C$ and $u \in \mathbb{I}$ the following conditions on $c \in C$ and $u \in \mathbb{I}$ are equivalent:*

$$\begin{array}{l} \alpha: \ \exists_\alpha \left[\neg\neg(c^{\alpha}=\bot) \ \& \ (u^{\alpha}=\top)\right] \\ \beta: \ \exists_\beta \left[\neg\neg(c^{\beta}=\bot) \ \& \ (u^{\beta}\neq\bot)\right] \\ \gamma: \ \exists_{\gamma,v} \left[(c^{\gamma}\neq\top) \ \& \ (v<u) \ \& \ (v^{\gamma}=\top)\right] \end{array}$$

(Note that the γ-condition will tend to be much more computationally feasible than the other two.) We need three implications:

$\alpha \Rightarrow \gamma$:

Given α let v be the unique element in $\mathbb{I}$ such that $v^{\alpha} = \odot$ and γ the result of following α with a $\bot$-zooming (and use $(\neg\neg(c^{\alpha}=\bot) \Rightarrow \neg\neg(c^{\gamma}=\bot) \Rightarrow (c^{\gamma}\neq\top)$.)

$\gamma \Rightarrow \beta$:

Given γ and v, we may assume that γ is the minimal zooming sequence for the task. We know that it is non-empty since $v < \top$ and $v^{\gamma} = \top$. If γ ends with an $\top$-zooming then the sequence obtained by removing that final $\top$-zooming would work as well. Thus from its minimality we may infer that γ ends with a $\bot$-zooming. Let γ' be the zooming sequence obtained by removing that final $\bot$-zooming. Since $v^{\gamma'\vee}=\top$ we know that $v^{\gamma'} \geqq \odot$, hence $u^{\gamma'} > \odot$ and $u^{\gamma'\wedge} \neq \bot$. Since $c^{\gamma'\vee} \neq \top$ the continuous zooming condition says that $\neg\neg[c^{\gamma'\wedge} = \bot]$. Thus we finish by defining β to be the result of following γ' with a $\top$-zooming.

$\beta \Rightarrow \alpha$:

Define α to be the result of following β with a sufficient number of $\bot$-zoomings to insure $u^{\alpha} = \top$.

Now for the Dedekind-cut conditions.

$U(c)$ is an open updeal:

Suppose $u \in U(c)$. Let γ, v be such that $c^{\gamma} \neq \top$, $v < u$ and $v^{\gamma} = \top$. Then for any $u' > v$ we have the same three conditions with u' instead of u, hence $u' \in U(c)$ for all $u' > v$.

$L(c)$ and $U(c)$ almost cover:

Given $\ell < u$ choose $\ell < k < v < u$: [147] Let γ be a zooming sequence (say the shortest one) such that $k^{\gamma} = \bot$ and $v^{\gamma} = \top$. The separation condition on C says that either $c^{\gamma} \neq \top$ or $c^{\gamma} \neq \bot$. In the first case we have $u \in U(c)$ and, dually, in the second case $\ell \in L(c)$.

$L(c)$ and $U(c)$ disjoint:

We wish to reach a contradiction from the assumption that there is $c \in A$, $q \in \mathbb{I}$ and zooming sequences σ, τ such that $\neg\neg(c^{\sigma} = \bot)$, $q^{\sigma} = \top$, $\neg\neg(c^{\tau} = \top)$ and $q^{\tau} = \bot$. We will

---

[146] $\neg\neg[x=\bot] \Rightarrow \neg[x\neq\bot] \Rightarrow \neg[x \neq \bot] \wedge \Big[[x\neq\bot]\vee[x\neq\top]\Big] \Rightarrow \Big[[\neg[x\neq\bot]\wedge[x\neq\bot]\Big] \vee \Big[\neg[x\neq\bot]\wedge[x\neq\top]\Big] \Rightarrow \Big[\neg[x\neq\bot]\wedge[x\neq\top]\Big] \Rightarrow [x\neq\top]$

[147] For example, $k' = \ell|u$, $v = k|u$.

settle for a weaker condition: $\neg\neg(c^{\sigma} = \bot)$ implies $c^{\sigma} \neq \top$ and $\neg\neg(c^{\tau} = \top)$ implies $c^{\tau} \neq \bot$. That is, we will reach a contradiction just from $c^{\sigma} \neq \top$, $q^{\sigma} = \top$, $c^{\tau} \neq \bot$ and $q^{\tau} = \bot$.

If σ were empty then $q = \top$ and it would not be possible for $q^{\tau} = \bot$. Dually, τ is non-empty. If σ were to start with $\top$-zooming we would know that $\hat{q} \neq \bot$ forcing $\overset{\vee}{q} = \top$ (the discrete coalgebra condition holds in $\mathbb{I}$) and thus $q^{\tau} = \top$, contradicting $q^{\tau} = \bot$. Hence σ starts with $\bot$-zooming and, dually, τ with $\top$-zooming. From $c^{\sigma} \neq \top$ we may infer $\overset{\vee}{c} \neq \top$ and from $c^{\tau} \neq \bot$ we infer $\hat{c} \neq \bot$. But the conjunctions of these two $\neq$s is precisely what the continuous coalgebra condition says can not happen.

We must show that this assignment of Dedekind cuts preserves the coalgebra structure. There is no difficulty in showing that $\langle L(\bot), U(\bot)\rangle$ and $\langle L(\top), U(\top)\rangle$ are what they should be. What we must show is $\langle L(\hat{c}), U(\hat{c})\rangle = \overset{\wedge}{\langle L(c), U(c)\rangle}$ (the other equation, of course, is dual). Restated: we must show

$$\overset{\wedge}{\ell} \in L(\hat{c}) \quad \text{iff} \quad \odot \leq \ell \in L(c)$$

$$\hat{u} \in U(\hat{c}) \quad \text{iff} \quad \odot < u \in U(c)$$

The forward directions are immediate: if α is a zooming sequence such that $\neg\neg(\hat{c}^{\alpha} = \top)$ and $\overset{\wedge}{\ell}^{\alpha} = \bot$ then if α' is the result of following an $\top$-zooming with α we have $\neg\neg(c^{\alpha'} = \top)$ and $\ell^{\alpha'} = \bot$ (and if $\ell < \odot$ then replace it with $\odot$). The same argument works when $\hat{u} \in U(\hat{c})$ (and since $\hat{u}^{\alpha} = \top$ we know that $u > \odot$).

For the reverse direction, suppose $\ell \geq \odot$, $\neg\neg(c^{\alpha} = \top)$ and $\ell^{\alpha} = \bot$. Then α is necessarily non-empty (that is, $\ell \neq \bot$) and it can not start with a $\bot$-zooming (since $\ell \geq \odot$ implies $\overset{\vee}{\ell} = \top$). Let α' be the rest of the sequence after the initial $\top$-zooming. Then $\neg\neg(\hat{c}^{\alpha'} = \top)$ and $\overset{\wedge}{\ell}^{\alpha'} = \bot$ forcing $\overset{\wedge}{\ell} \in L(\hat{c})$

Finally, suppose $u > \odot$, $\neg\neg(c^{\alpha} = \bot)$ and $u^{\alpha} = \top$. If α is empty then the empty sequence also establishes $\hat{u} \in U(\hat{c})$. If α starts with a $\top$-zooming we use the same sort of argument just above to establish $\hat{u} \in U(\hat{c})$. If α starts with a $\bot$-zooming then for any $u > \odot$ it is the case that $u^{\alpha} = \top$, hence we need to show that everything in $\mathbb{I}$ other than $\bot$ is in $U(\hat{a})$. But we may infer $\neg\neg(c^{\alpha} = \bot) \Rightarrow (c^{\alpha} \neq \bot) \Rightarrow (\overset{\vee}{c} \neq \bot)$ and the continuous coalgebra condition says that $\neg\neg(\hat{c} = \top)$. For β the empty sequence we thus have $\neg\neg(\hat{c}^{\beta} = \top)$ and $\hat{u}^{\beta} \neq \bot$ forcing $\hat{u} \in U(\hat{c})$.

32. *Appendix:* The Peneproximate Origins

I always disliked analysis. Algebra, geometry, topology, even formal logic, they captivated me; analysis was different.

My attitude, alas, wasn't improved when I was supposed to tell a class of Princeton freshmen about numerical integration. I was expected to tell them that trapezoids were better than Riemann and Simpson was better than trapezoids. I was not expected to prove any of this.

I was appalled by the gap between applied mathematical experience and what we could even imagine proving. How does one integrate over all continuous functions to arrive at the expected error of a particular method?

Of course one can carve out finite dimensional vector spaces of continuous functions and compute an expected error thereon. But all continuous functions? It's easy to prove that there is no measure—not even a finitely additive measure—on the set of all continuous functions assuming at least that we ask for even a few of the most innocuous of invariance properties. Yet experience said that there was, indeed, such a measure on the set of functions one actually encounters.

But it wasn't just a problem in mathematics: I learned from physicists that they succeed in coming to verifiable conclusions by pretending to integrate over the set of all paths between two points. Again it is not hard to prove that no such "Feynman integral" is possible once one insists on a few invariance properties.

Even later I learned (from the work of David Mumford) about "Bayesian vision": in this case one wants to integrate over all possible "scenes" in order to deduce the most probable interpretations of what is being seen. A scene is taken to be a function from, say, a square to shades of gray. It would be a mistake to restrict to continuous functions—sharp contour boundaries surely want to exist. Quite remarkable "robotic vision" machines had been constructed for specific purposes by judiciously cutting down to appropriate finite-dimensional vector spaces of scenes. But once again, there is no possible measure on sets of all scenes which enjoy even the simplest of invariance conditions.

Thus three examples coming with quite disparate origins—math, science, engineering—were shouting that we need a new approach to measure theory.

One line of hope arose from the observation that the non-existence proofs all require a very classical foundation. There's the enticing possibility that a more computationally realistic setting—as offered, say, by "effective topoi"—could resolve the difficulties. A wonderful dream presents itself: the role of foundations in mathematics—and its applications—could undergo a transformation similar to the last two centuries' transformation of geometry.

Geometry moved from fixed rigidity to remarkable flexibility and—in the last century—that liberal view became a critical tool in physics. We learned that there was a trade-off between physics and geometry; we could still insist on classical (Euclidean) geometry but only at the expense of a cumbersome physics. We no longer even view most questions on the nature of geometry to be well put unless first the nature of physics be stipulated and—of course—vice versa.

Could we now learn the same about foundations? Elementary topoi provide a general setting for shifting foundations reminiscent of the role of Riemannian manifolds in geometry. Might the trade-off between physics and geometry be replicated for physics and foundations? Two hundred years ago there was only one geometry. It was more than taken for granted; it was deemed to be certain knowledge.

Of course the geometry we now call Euclidean was certain; it may not be innate but it is inevitable. I have no doubt that if we lived in a universe with a visibly non-zero curvature we would get around to building our blackboards (or whatever we teach calculus with) with zero curvature.[148]

It must be deemed remarkable that we learned to think—and make correct predictions—in non-Euclidean geometry. We learned to imagine living in a 3-sphere, in spaces of higher genus, even in projective space. The representation theorems for Riemannian manifolds (long before they were all proved) played a critical role in that process; and so it is with foundations. Bill

[148] In this neighborhood, of course, Euclidean geometry is the natural geometry from the very beginning. When I was a kid a friend had measured the distance around a giant tree. We estimated the tree's width by solving the same problem on a little fruit-juice glass. We never questioned that the same ratio would hold for giant trees and little fruit-juice glasses.

Lawvere taught us that with a few topoi on hand for comparison we can learn to shift our foundations between what's called classical and what's called (alas) intuitionistic. Again, the representation theorems play a critical role: in a fully classical setting a category of sheaves on a space can support a fully intuitionistic logic—change the topology and you can revert to the classical.

Coming back to earth: I must confess that the perfectly obvious idea that one should first establish ordinary integration in the right way on something as simple as the closed interval, that simple idea took longer than it should have (it had to await a day's boat trip in Alaska, of all places). For some years I preached this doctrine to the category/topos crowd and some trace of those preachings can be found scattered in the literature.[149] In September 1999 at an invited talk at the annual CTCS meeting (held that year in Edinburgh) I even characterized the mean value of real-valued continuous functions on the closed interval as an order-preserving linear operation that did the right thing to constants and had the property that the mean value on the entire interval equaled the midpoint of the mean-values on the two half intervals. I described it with a diagram that used (twice) the canonical equivalence between $\mathtt{I}$ and $\mathtt{I} \vee \mathtt{I}$.

But one equivalence, even used twice, doesn't bring forth the general notion: it doesn't prompt one to invent ordered wedges; without ordered wedges one doesn't define zoom operators nor discover the theorem on the existence of standard models (Theorem 10.5, p28). One doesn't learn how remarkably algebraic real analysis can become.

What I needed was someone to kick me into coalgebra mode. Three months later two guys did just that and on the 22nd day of December I wrote to the category list:

> There's a nice paper by Duško Pavlović and Vaughan Pratt. It's entitled
> On Coalgebra of Real Numbers [150] and it has turned me on.

A solution, alas, for the three motivating problems still awaits; but, at least, now I like analysis.

33. *Addendum:* A Few Latex Macros

$\mathtt{I}$

`\CI:` `\scalebox{1.14}{\ensuremath{\tt I}}`
Requires graphics package. In its absence use the ossia, `{\ensuremath{\tt I}}` .
There are those who insist that this portrays a copulation of $\top$ and $\bot$. (Indeed, they go on to say that about the entire subject.)

$\bot \quad \top$

`\bt:` `\scalebox{.83}{\ensuremath{\bot}}`
`\tp:` `\scalebox{.8}{\ensuremath{\top}}`
Ossias: `{\ensuremath{\bot}}` and `{\ensuremath{\top}}`

[149] *e.g.,* Abbas Edalat and Martín Hötzel Escardó, Integration in real PCF, LICS 1996, Part I (New Brunswick, NJ). *Information and Computation* 160 (2000), no. 1–2, p128–166.
[150] Later published as The continuum as a final coalgebra CMCS'99 Coalgebraic methods in computer science (Amsterdam, 1999). *Theoret. Comput. Sci.* 280 (2002), no. 1-2, p105–122.

$$⨏$$

`\Fint:  \;{\rotatebox{103}{\scalebox{.55}{$\int$}}}\hspace{-4.23mm}\int`
The main "`\int`" is elayed until the end so that as a macro it accepts sub- and superscripts. A less than satisfactory ossia: `\;-\hspace{-4.5mm}\int`

[1] [12] [123]

`\fna[1]: {\ensuremath{^[\footnote{\hs{-6.5}$^[\hs{3.75}^]$\hs4#1}^]}}`
`\fnb[1]: {\ensuremath{^[\footnote{\hs{-10.25}$^[\hs{7.75}^]$\hs4#1}^]}}`
`\fnc[1]: {\ensuremath{^[\footnote{\hs{-14}$^[\hs{11.75}^]$\hs4#1}^]}}`
(`\hs{`x`}` means `\hspace{`x`pt}`.) Its *raison d'être* is for use at the end of math display lines (and should, in that case, usually be preceded with `\;\;`). Do not use for an asterisk at the end of the title line. Other delimiters, of course, could be used: $^{\{1\}}$ $^{\langle 2\rangle}$ $^{|3|}$ $^{\lfloor 4\rceil}$ $^{\lceil 5\rfloor}$
Ossia: `\fn[1]: {\ensuremath{^[\footnote{\,#1}^]}}`

$$\hat{x} \quad \overset{\vee}{x} \quad \overset{\leftrightarrow}{x}$$

`\tz[1]: \stackrel{\wedge}{#1}`   `\bz[1]: \stackrel{\vee}{#1}`
`\mz[1]: \stackrel{\hspace{.1}\scalebox{.7}{$\leftrightarrow$}}{#1}`
Use `\hat x` and `\check x` in subscripts: $\hat{x}$ $\check{x}$.

$$\overset{\mathbf{.}}{u} \qquad (x+y)^{\bullet}$$

`\dt[1]: \stackrel{\mbox{\bf.}}{#1}`
`\hdot:  \begin{picture}(0,8)\put(-1,8.1){\bf.}\end{picture}`
Use `\dot x` in subscripts: $\dot{x}$. In footnotes use `\put(-1,6)` instead of `\put(-2,8.1)`.

$$\overset{\wedge}{|} \quad \overset{\vee}{|} \quad |$$

`\mtz:  scalebox{.7}[.9]{\begin{picture}(9,15)\put(2.4,0){$|$}`
`       \put(1,9){\scalebox{.8}{$\wedge$}}\end{picture}}`
`\mbz:  \scalebox{.7}[.9]{\begin{picture}(9,15)\put(2.4,0){$|$}`
`       \put(.8,9.1){\scalebox{.8}{$\vee$}}\end{picture}}`
In footnotes:
`\ftz:  \scalebox{.7}[.9]{\begin{picture}(9,11)\put(2.4,0){$|$}`
`       \put(1.05,6){\scalebox{.9}{$\wedge$}}\end{picture}}}`
`\fbz:  \scalebox{.7}[.9]{\begin{picture}(9,11)\put(2.4,0){$|$}`
`       \put(1.05,6.5){\scalebox{.9}{$\vee$}}\end{picture}}}`
Ossias: `\!\stackrel{\wedge}{|}\!` and `\!\stackrel{\vee}{|}\!`
`\md:  \mbox{\huge${\mbox{\Large$|$}}$}` In footnotes use `large` not `Large`.
`\wmd:  \;\md\;` Used in formulas such as $((\perp|\top)\big|(\perp|\odot))\ \big|\ (\perp|x)$

$$\mathbb{I} \quad \mathring{\mathbb{I}}$$

```
\IC:    \ensuremath{\mathbb I}
\ic:    \begin{picture}(9,12)\put(0,0){$\IC$}\put(.2,9)
        {\scalebox{.7}{$\circ$}}\end{picture}
```
In footnotes use:
```
\ics    \begin{picture}(6,9)\put(0,0){$\IC$}\put(.15,6.5)
        \scalebox{.7}{$\circ$}}\end{picture}}
```

$$x \leqq y \qquad x \geqq y \qquad x \lesseqgtr y \qquad x \gtreqless y$$

```
\lq     \begin{picture}(17,0)\put(3.65,0){$\leq$}\put(4,-1.9)
        {$\color{white}\rule{8.5pt}{5.8pt}$}\put(4,0){$=$}\end{picture}}
```
For `\gq` replace `(3.65,0)` with `(4,0)` and `\leq` with `\geq`
```
\lg:    \begin{picture}(13,0)\put(1.7,2.9){$<$}
        \put(2,-2.4){$>$}\end{picture}
```
For `\gl` swap `>` and `<`

$$x^{-1}y^{-m}z^{-2}$$

```
\i:     \def\i{\inv}\newcommand{\inv}[1]
        {^{{\scalebox{1.3}[.76]{-}}\hs{-.5}#1}}
```
Use `x\i1y\i mz\i2` instead of `x^{-1}y^{-m}z^{-2}` $(x^{-1}y^{-m}z^{-2})$

$$⅋$$

```
\Par:  \begin{picture}(14,0)\put(2,7.5){\scalebox{-.9}{\&}}\end{picture}
```
Ossia: complain to Jean-Yves.

$$x \multimap y \qquad x \circ\!\!\multimap y \qquad x \overset{\bullet}{\circ\!\!\multimap} y$$

```
\loli:  \begin{picture}(20,0)\put(3,0){$-$}\put(4,0)
        {$-$}\put(11.4,0){$\circ$}\end{picture}
```
In footnotes use `\put(10,0)` instead of `\put(11.4,0)` : $\multimap$
```
\bimp:  \begin{picture}(23,0)\put(2,0){$\circ$}
        \put(3,0){$\loli$}\end{picture}
\abs:   \stackrel{^\bullet}{\bimp}
```

$$\circ + - \circ \circ$$

```
\z:     \scalebox{1.3}{\begin{picture}(6,0)
        \put(0,-.6){$\circ$}\end{picture}}
```
For signed-binary and “symmetric ternary” expansions

↺ ↻

```
\lloop[1]:    \scalebox{1.5}{\rotatebox{#1}{\begin{picture}(0,0)\put(0,0)
      {\circle{16}}\put(-8.32,0){$\color{white}\rule{8.32pt}{8.3pt}$}
      \put(-1.3,6.9){\rotatebox{-13}{\scalebox{.5}{$<$}}}\end{picture}}}
```

For the curlyheaded variation (as on page 85) replace last line with:

```
      \put(-2.5,6.9){\rotatebox{-9}{\scalebox{.5}{$\prec$}}}\end{picture}}}
\rloop[1]:    \reflectbox{\lloop{#1}}
```

$$[\![w]\!] \quad (\!(x)\!) \quad \langle\!\langle y\rangle\!\rangle \quad \lfloor z\rfloor$$

```
\banana[1]:   \,[\![#1]\!\!\,]
\pr[1]:       \,(\!(#1\,)\!\!\,)
\Ang[1]:      \langle\!\langle#1\rangle\!\rangle
\tr[1]:       \lfloor#1\rfloor
```

(and don't forget `\|v\|` for $\|v\|$)

◡̇ ◠̇

```
\fq:    begin{picture}(9,0)\put(0,0){$\ct$}\put(0,3.75)
        {\color{white}\rule{9pt}{5pt}}\end{picture}}
```

For `\tq` replace 3.75 with -2.85. Only with great reluctance did I forgo using the standard calendar lunar symbols for first- and third-quarter. Ossias: `\mbox{\small Q}_1` and `\mbox{\small Q}_3`

$\succ$

```
\venturi:     \scalebox{.8}{\begin{picture}(17,7)\qbezier(5,5.5)(6,4)(12,4)
              \qbezier(5,.5)(6,2)(12,2)\end{picture}}
```

34. *Addendum:* Heyting Scales [2009–04–05]

In [94] (p44) we pointed out that the support operator of a chromatic scale may be defined using the Heyting structure (which was defined using the chromatic structure). But that does not establish that any scale with a Heyting structure is, in fact, a chromatic scale. Hence this addendum.

In fact, we use less than the full Heyting structure. For the chromatic structure it suffices that a scale be a "negated scale" and that is precisely what allows us to show that an order-complete scale is chromatic (see below).

Let's establish some definitions.

The quickest definition of a **Heyting semi-lattice** (for a category theorist) is a poset which when viewed as a category is exponential (sometimes "cartesian closed"). That is, a meet semi-lattice with top and a binary operation whose values are denoted $y \to z$ characterized by the "adjointness condition:"

$$x \leq y \to z \text{ iff } x \wedge y \leq z$$

Put another way, $y \to z$ is the largest element whose meet with y is bounded by z. [151]

[151] The adjointness condition is equivalent to the three equations:

$$\begin{aligned} \top \to z &= z \\ (y \wedge z) \to z &= \top \\ x \wedge (y \to z) &= x \wedge ((x \wedge y) \to z) \end{aligned}$$

A **Heyting algebra** is a lattice with bottom that is a Heyting semi-lattice. Any linearly ordered set with top and bottom is such:

$$u \to v \;=\; \begin{cases} \top & \text{if } \; u \leqq v \\ v & \text{if } \; v < u \end{cases}$$

Clearly then, in any linear chromatic scale $u \to v$ is constructable as $(u \multimap v) \vee v$. The linear representation theorem for chromatic scales tells us that the adjointness condition that characterizes the Heyting arrow holds for all chromatic scales (as do all universal Horn sentences).

It should be noted, though, that the Heyting algebras that so appear are rather special. They all satisfy the "equation of linearity" $(u \to v) \vee (v \to u) = \top$. There are no further equations, or for that matter, universally quantified first-order properties: any countable Heyting algebra that satisfies the equation of linearity may be faithfully represented in a power of $\mathbb{I}$ (it's not hard to show that SDI Heyting algebras enjoy the disjunction property).

The lemma that any scale with a Heyting structure is a chromatic scale uses less than the entire Heyting structure: we need only the arrow operation when targeted at the bottom, $(x \to \bot)$. So:

Define a **negated semi-lattice** to be a meet semi-lattice with top and bottom and a unary operation, whose values are denoted $\neg y$, that delivers the largest element disjoint from y. That is, it satisfies the adjointness condition:

$$x \leqq \neg y \quad \text{iff} \quad x \wedge y = \bot \qquad {}^{[152]}$$

Note that the characterization of $\neg x$ as the largest element disjoint from x easily implies that negation is contravariant, hence double negation is covariant. Double negation is inflationary: $x \leqq \neg\neg x$ (because $x \wedge \neg x = \bot$). The contravariance of negation then says $\neg\neg\neg x \leqq \neg x$. But a special case of $x \leqq \neg\neg x$ is $\neg x \leqq \neg\neg\neg x$ (it's the case obtained by replacing x with $\neg x$). Thus $\neg x = \neg\neg\neg x$. In particular double-negation is a closure operation (inflationary and idempotent)

A consequence is:

34.1 LEMMA: *In negated semi-lattices* $x \wedge y = \bot$ *iff* $x \wedge \neg\neg y = \bot$.

Because $x \wedge y = \bot$ iff $x \leqq \neg y$ iff $x \leqq \neg(\neg\neg y)$ iff $x \wedge \neg\neg y = \bot$.

An important equation for us is the Lawvere-Tierney condition on a closure operation:

34.2 LEMMA:

$$\neg\neg(x \wedge y) = \neg\neg x \wedge \neg\neg y$$

The verification of the equations is as follows: for the 1st equation note that $t \le \top \to z$ iff $t \wedge \top \le z$, that is, $t \le \top \to z$ iff $t \le z$; for the 2nd equation note that $\top \wedge (y \wedge z) \le z$ hence $\top \le ((y \wedge z) \to z)$; for the 3rd equation note that $t \le x \wedge (y \to z)$ iff $t \le x$ and $t \wedge y \le z$ whereas $t \le x \wedge ((x \wedge y) \to z)$ iff $t \le x$ and $t \wedge (x \wedge y) \le z$ and the two conditions are clearly the same.

For the derivation of the adjointness condition from the equations assume first that $x \le (y \to z)$. Then $x \wedge y \le y \wedge (y \to z) = y \wedge ((y \wedge \top) \to z) = y \wedge (\top \to z) = y \wedge z \le z$. Second, assume $x \wedge y \le z$ or, as it will appear below, $x \wedge y \wedge z = x \wedge y$. Then $x = x \wedge \top = x \wedge ((y \wedge z) \to z) = x \wedge ((x \wedge y \wedge z) \to z) = x \wedge ((x \wedge y) \to z) = x \wedge (y \to z) \le y \to z$.

For the independence of the three equations: taking $y \to z$ as $\top$ satisfies just the 2nd and 3rd equations; taking it as $\bot$ satisfies just the 1st and 3rd ; taking it as $y \multimap z$ in any non-trivial scale satisfies just the 1st and 2nd.

[152] Its equational characterization is given by:

$$\begin{aligned} \neg\bot &= \top \\ \neg\top &= \bot \\ x \wedge \neg y &= x \wedge \neg(x \wedge y) \end{aligned}$$

Both the adjointness condition and the equations are obtained, of course, just by replacing the variable z with $\bot$ in the equations for Heyting semi-lattices. The equivalence of the two definitions and the independence examples are easily obtained by following through with that replacement.

Because the covariance of double negations easily implies (indeed, is equivalent with) $\neg\neg(x \wedge y) \leqq \neg\neg x \wedge \neg\neg y$. For the other direction, $\neg\neg x \wedge \neg\neg y \leqq \neg\neg(x \wedge y)$, it suffices to show $\neg\neg x \wedge \neg\neg y \wedge \neg(x \wedge y) = \bot$. The last lemma (used twice) says that this last equation is equivalent to the obvious equation $(x \wedge y) \wedge \neg(x \wedge y) = \bot$.

Define a **negated scale** to be a scale with a unary operator satisfying the equations for a negated semi-lattice.

34.3 LEMMA: *In a negated scale, $\neg x$ is an extreme point.*

It suffices to show $(\neg x)^{\vee} \leqq \neg x$ and for that it suffices to show $(\neg x)^{\vee} \wedge x = \bot$. So: $(\neg x)^{\vee} \wedge x \leqq (\neg x)^{\vee} \wedge x^{\vee} = (\neg x \wedge x)^{\vee} = \bot^{\vee} = \bot$. Recall that the $\top$- and $\bot$-zooming operations have the same fixed points. In particular, $(\neg x)^{\wedge} = \neg x$.

We have now established:

34.4 LEMMA: *Double negation satisfies the four defining equations for the support operation:*

$$\begin{aligned} \neg\neg\bot &= \bot \\ (\neg\neg x)^{\wedge} &= \neg\neg x \\ x \wedge \neg\neg x &= x \\ \neg\neg(x \wedge y) &= \neg\neg x \wedge \neg\neg y \end{aligned}$$

In [94] (p44) we identified $\overline{x}$ not as the double negation but as $(\neg x)^{\bullet}$. Since $\neg x$ is an extreme point we know that it is complemented and its complement (we're in a distributive lattice recall) is clearly its maximal disjoint element.

The scale structure allows us to construct the Heyting arrow operation starting with negation: $u \to v = \neg(u \overset{\wedge}{|} v) \vee v$. It's easy to find negated lattices that aren't Heyting algebras, indeed aren't even distributive lattices: take any lattice with top and formally adjoin a bottom element (even if it already has one); the result is a negated lattice. (A semi-lattice with bottom, on the other hand, is a Heyting semi-lattice iff every principal filter is a negated semi-lattice.)

We close with

34.5 LEMMA: *All order-complete/injective scales are negated scales, hence chromatic.*

We saw in Section 24 (p53–57) that the scales in question are of the form $\mathcal{C}(X)$ where X is extremely disconnected. Given $f \in \mathcal{C}(X)$ let $C \subseteq X$ be the closure of $\{\, x \in X \,:\, f(x) > \bot \,\}$. Since C is a "clopen" it has a continuous characteristic function. It is easy to see that χ_C works as $\neg f$. [153]

[153] For an example of a metrically complete chromatic scale that is not order-complete (hence not injective) let $\Pi_J \mathbf{I} = \mathbf{I}^J$ be an uncountable cartesian power of $\mathbf{I}$ and $S \subset \mathbf{I}^J$ the subscale of elements that with the exception of a countable subset of J are equal to an element in the image of the diagonal map (to wit, the unique scale homomorphism from $\mathbf{I}$ to $\mathbf{I}^J$). The chromatic structure on S is clear as is the fact that any order-complete scale that contains it also contains a copy of $\mathbf{I}^J$. (Chromatic complete scales thus need not be order-complete but it is fairly easy to show that they do have countable sups and infs.)

35. *Addendum:* Wilson Space [2009–04–25] In the 1999 category-list post in which I first described the final-coalgebra characterization of $\mathbf{I}$ the ordered-wedge functor was defined as follows:

> In the category of posets with top and bottom consider the binary functor, $X \vee Y$, obtained by starting with the disjoint union $X;Y$, with everything in X ordered below Y,[154] and then identifying the top of X with the bottom of Y.

That posting had a PS:

> Just for comparison, consider the category of posets and the functor that sends X to $X;1;X$. The open interval is an invariant object for this functor but it is not the final co-algebra. For that we need—as we called it in *Cats & Alligators*—Wilson space. Actually, not the space but the linearly ordered set, most easily defined as the lexicographically ordered subset of sequences with values in $\{-1, 0, 1\}$ consisting of all those sequences such that $a(n) = 0 \Rightarrow a(n+1) = 0$ for all n (take a finite word on $\{-1, 1\}$ and pad it out to an infinite sequence by tacking on 0s).

Actually, in *Cats & Alligators*[155] Wilson space was not viewed as a poset but a topological space. So let's work in the category of spaces (we'll come back to the poset view later). We topologize $X; 1; X$ by taking it as $\boldsymbol{Scone(X+X)}$ the **scone** of the disjoint union of two copies of X. The scone of a space is the space that results when a new point is adjoined whose only neighborhood is the (resulting) entire space. (Restated: a subset of $Scone(X)$ is open iff it is entire or an open subset of X.) The new point is called the **focal point** (all sequences converge to it). $Scone(X)$ classifies continuous partial maps with open domains: continuous maps $Y \to Scone(X)$ are in natural correspondence with continuous maps $U \to X$ where U is an open subset of Y. The final coalgebra of $Scone(-)$ is the space of "extended natural numbers," that is, the natural numbers plus a point at infinity, topologized by taking as its only open nonempty subsets the infinite updeals (there's only one nonempty finite updeal, to wit, the one-element set $\{\infty\}$). The focal point is 0. Given a space X and a continuous partial map with open domain $f : X \to X$, the induced map from X to the extended natural numbers sends $x \in X$ to $\sup\{\ n : f^n(x)\downarrow\ \}$ (using the computer-science convention that $\downarrow$ means "the expression to the left is actually defined"). The final coalgebra, of course, needs a coalgebra structure; it is given by the predecessor map (where it is understood that ∞ is a fixed-point and that the predecessor of 0 is undefined). The map just described from X to the extended natural numbers is then the unique co-homomorphism between coalgebras.

I find it remarkable that the final coalgebra for the functor $Scone(X+X)$ is the same space that we defined for an entirely different reason in *Cats & Alligators.*[156].

[154] I was borrowing the computer-science use of the semi-colon for joining a pair of imperatives. Am I right in believing that this first appeared in Kemeny's BASIC?

[155] Freyd, Peter and Scedrov, Andre, Categories, Allegories, *North-Holland Publishing Co., Amsterdam*, 1990

[156] It arose in the proof of the "geometric representation theorem" for intuitionistic logic. After establishing a completeness theorem for intuitionistic logic in the semantics arising in set-valued functor categories we wish to establish a completeness theorem for the category of sheaves on the reals. If $T : \mathcal{B} \to \mathcal{A}$ has the property (reminiscent of covering maps in topology) that for any $f \in \mathcal{A}$ there is not only $g \in \mathcal{B}$ such that $T(g) = f$ but for every $B \in \mathcal{B}$ such that $T(B) = \mathrm{Dom}(f)$ there is $g \in \mathcal{B}$ such that $\mathrm{Dom}(g) = B$ and $T(g) = f$ then the functor induced by composition $\mathcal{S}^T : \mathcal{S}^{\mathcal{A}} \to \mathcal{S}^{\mathcal{B}}$ faithfully preserves the semantics of first-order logic of the two categories. For any $\mathcal{A}$ let $\mathcal{P}$ be the "path tree," to wit the partially ordered set of non-empty finite sequences of composable maps in $\mathcal{A}$ ordered by prolongation on the right. Any partially ordered set may be viewed as a category. and the obvious "forgetful functor" from $\mathcal{P}$ to $\mathcal{A}$ is an example of a functor that induces a functor between functor categories that faithfully preserves the semantics of first-order logic. Each connected component of $\mathcal{P}$ is a rooted tree and $\mathcal{S}^{\mathcal{P}}$ is a cartesian product of functor categories based on rooted trees. Any countable tree may be covered with an ever-bifurcating tree.

As can any category of set-valued functors on a poset, the category of set-valued functors from an ever-bifurcating tree may be viewed as the category of sheaves on a space, to wit, the ever-bifurcating tree where the open sets are all the ever-bifurcating subtrees. When any space is made sober (see the first sentence of [59], p25) the category of sheaves remains the same. Wilson

I'll call a coalgebra on the functor whose values are denoted $Scone(X+X)$ a **partial-interval coalgebra** and continue to call its final coalgebra **Wilson space** and denote it as $\mathbb{W}$. A partial-interval structure on X is given by a pair of continuous partial self-maps with disjoint open domains. I'll denote the values of the partial maps as $\hat{x}$ and $\overset{\vee}{x}$ and call them **partial** $\top$- and **partial** $\bot$**-zooming**.

We construct $\mathbb{W}$ as the disjoint union $\{-,+\}^{\star} \cup \{-,+\}^{\mathbb{N}}$, to wit, the finite and infinite words on a two-element alphabet $\{-,+\}$. The word "path" has been used to encompass both finite paths (words) and infinite paths (sequences). It has two orderings of interest to us. Besides the total ordering from 1999 there's a partial ordering defined by prolongation on the right. With this ordering the finite paths form a tree with the empty word as root. (I find it easiest to view the tree not as going up or down but sideways. The root is on the left.) For each finite word we obtain a subtree with that finite path as root. We topologize $\mathbb{W}$ by taking all such rooted subtrees as basic open subsets. It may be noticed that the subspace of infinite paths is none other than Cantor space.

The partial-interval structure on $\mathbb{W}$ is such that $\hat{p}\downarrow$ iff the **head** of p is $+$ in which case $\hat{p}$ is its **tail** (the head is the first element in the path, the tail is what's left after the head is removed). $\overset{\vee}{p}$ is defined dually. Note that neither zooming is defined on the empty word (it doesn't have a head).

Given a coalgebra structure on a space X we construct $f : X \to \mathbb{W}$ by defining $f(x) \in \mathbb{W}$ as the path, finite or infinite, obtained by iterating the parallel procedure:

$$\left[\begin{array}{l}\texttt{If } \hat{x}\downarrow \texttt{ then emit } \text{“}+\text{”}\\ \texttt{and replace } x \texttt{ with } \hat{x}\end{array}\right] \;\Big|\Big|\; \left[\begin{array}{l}\texttt{If } \overset{\vee}{x}\downarrow \texttt{ then emit } \text{“}-\text{”}\\ \texttt{and replace } x \texttt{ with } \overset{\vee}{x}\end{array}\right]$$

It is transparent that such defines a coalgebra homomorphism. For its continuity note that the inverse image of a basic open set rooted at a given finite path is the domain of the partial map determined by composing the sequence of $\top$- and $\bot$-zoomings corresponding to the $+$s and $-$s in the given path.

In *Cats & Alligators* we needed an open continuous map from $\mathbb{I}$ to $\mathbb{W}$ (a "Freyd curve"). We can construct the unique coalgebra map for a particular partial interval structure defined as follows. The domain of $\top$-zooming will be the strictly positive elements in $\mathbb{I}$. The map from $(0,1]$ to $[-1,1]$ will be piecewise affine with infinitely many pieces. $\{2/3^n\}_{n\in\mathbb{N}}$ is the sequence of critical points. The critical values alternate between $+1$ and -1 with $\overset{\wedge}{1}= 1$. As previously $\bot$-zooming is defined as $-\widehat{(-x)}$.

The resulting $\mathbb{I} \to \mathbb{W}$ can be (and originally was) defined using "symmetric ternary expansions." Not signed-binary expansions. Every element in $\mathbb{I}$ may be described as

$$2\sum_{n=1}^{\infty}\frac{a_n}{3^n} \qquad [157]$$

space is none other than the ever-bifurcating tree made sober.

In order to move the semantics to a space more familiar than $\mathbb{W}$ we need an open continuous map from a familiar space to $\mathbb{W}$. If $X \to Y$ is open, continuous and onto then the induced map $Sh(Y) \to Sh(X)$ faithfully preserves the first-order semantics. We want an open continuous map onto $\mathbb{I} \to \mathbb{W}$. (All sorts of familiar spaces have open continuous maps onto $\mathbb{I}$. As just one example to get started: the sine function from $\mathbb{R}$ to $\mathbb{I}$.) Wonderfully enough the function we described in *Cats & Alligators*, the one that had earned the name "The Freyd Curve" in the 70s, is the induced map $\mathbb{I} \to Scone(\mathbb{W})$ arising from a particular coalgebra structure on $\mathbb{I}$.

[157] In *Cats & Alligators* the 2 was omitted and we worked in the interval $[-\frac{1}{2},+\frac{1}{2}]$ (which, by the way, made the forthcoming characterization of the elements with non-unique expansions a tiny bit harder).

where $a_n \in \{-1, 0, +1\}$. As in Section 30 (p66–71) I'll suppress the 1s and use the symbols "$-\circ+$." The non-integers in $\mathtt{I}$ of the form $(2m+1)/3^n$ have two such expansions, all other elements have just one.

The induced function $\{-1, 0, +1\}^{\mathbb{N}} \to \mathbb{W}$ can be described with a four-state automaton. E is the initial state (actually, all choices of initial state work):

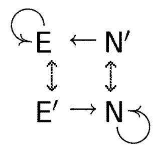
Next State

	E	E′	N	N′
+	E	N	N	E
∘	E′	E	N′	N
−	E	N	N	E

Output

	E	E′	N	N′
+	+	+	−	−
∘				
−	−	−	+	+

The ∘ input toggles the pairs E, E′ and N, N′. The two non-∘ inputs have the same next-state behavior: E and N are stationary for non-∘ inputs and they are the only non-∘-input targets. In the diagram below, the vertical gray (double) arrows show the next-state behavior for the ∘ input, the horizontal and circular arrows show the next-state behavior for non-∘ input:



As for output: when z is the input there is no output (a stammer); non-∘ input is echoed in the two left-hand states, E, E′ and negated in the two right-hand states, N, N′.

The first task is to show that the resulting function is defined not just on sequences but on elements of the standard interval, that is, the same output is engendered by sequences that name the same interval element. We need to consider the output of two machines starting in the same place but one being fed the sequence $+----\cdots$ and the other $\circ++++\cdots$. We'll do better with a single machine but with two demons jumping from state to state each according to the commands issued by its appointed sequence. We start them at the same state. For each of those four possible states the jump commanded by the input $+$ produces the same output as that commanded by $\circ+$ but the demons will land in different states. One will be in E the other N and there they'll henceforth remain. One of the demons will echo the input, the other will negate it, which is just what is needed for the output engendered by a constant sequence of +s for one demon to be the same as that engendered by a constant sequence of −s for the other.

The "Freyd curve" in *Cats & Alligators* was different but can also be described with a four-state automaton:

Next State

	E′	E	N	N′
+	E	N	E	N
∘	E	E′	N′	N
−	E	N	E	N

Output

	E′	E	N	N′
+	+	+	−	−
∘				
−	−	−	+	+

Any one of the states may be taken as initial. (To get the function described in *Cats & Alligators*[158] start at E) The next state is always an adjacent state. The ∘ input toggles

[158] It was described there as the function that sends $2\sum_{n=1}^{\infty} a_n 3^{-n} \in \mathtt{I}$ to $\{(-1)^{n_i-1} a_{n_i}\}_{i=1}^{L} \in \mathbb{W}$ where $\{a_{n_i}\}_{i=1}^{L}$ is the result of removing all 0s from $\{a_n\}_{n=1}^{\infty}$ (and n_0 is understood to be 0).

the pairs E', E and N, N'. The non-$\circ$ inputs have the same next-state behavior: they both always target one of the two middle positions which fact together with the fact that all next states are adjacent determines their action. In the diagram below the gray (double) arrows show the next-state behavior for the $\circ$ input.

$$\mathsf{E}' \overset{\leftrightarrow}{\rightarrow} \mathsf{E} \leftrightarrow \mathsf{N} \overset{\leftrightarrow}{\leftarrow} \mathsf{N}'$$

As for output: when $\circ$ is the input there is no output (a stammer); non-$\circ$ input is echoed in the first two states, E', E and negated in the last two, N, N'.

That the resulting function is defined not just on sequences but on elements of the interval is obtained by an argument almost identical to that given for the first machine. We need, again, to consider the output of two demons starting at the same state, one behaving according to the sequence $+----\cdots$ and the other $\circ++++\cdots$. On each of the four possible states the jump commanded by the input $+$ produces the same output as that commanded by $\circ+$ but the demons will land in different states. One will be in E the other N and they'll spend the rest of eternity trading places: at each subsequent input one of the demons will echo, the other will negate, which is just what is needed for the output engendered by a constant sequence of $+$s for one demon to be the same as that engendered by a constant sequence of $-$s for the other.

For the continuity let's consider a more general setting. We are considering functions given by procedures that take finite paths of signals to finite paths in $\mathbb{W}$. If we take a very general view of procedure we are led to view functions of the form $m : A^\star \to B^\star$ where A and B are finite sets, which functions are covariant with respect to the prolongation ordering, that is, if $w, w' \in A$ and if w' is a prolongation of w then we require $m(w')$ to be a prolongation of $m(w)$ (which includes, of course, the possibility that $m(w') = m(w)$). Such a function induces a function $\tilde{m} : A^{\mathbb{N}} \to B^\star \cup B^{\mathbb{N}}$. In the classical case $A^{\mathbb{N}}$ and $B^{\mathbb{N}}$ are understood to name elements in an interval (and, traditionally, $A = B = \{0, 1, .., b-1\}$). In that case we need two conditions: first, for any infinite input path the induced output path is also infinite; second, if a pair of infinite input paths name the same element in the interval then so do the pair of infinite output paths. The resulting function on the interval is then automatically continuous. One would like, of course, that all continuous maps are so obtainable. Alas, even something as simple as multiplying by $3/4$ on the unit interval can not be so obtained when using ordinary binary expansions (if the input is the unique binary expansion for $2/3$ then no finite initial subpath has enough information to determine even the head of the intended output path). Signed-binary expansions were invented to take care of this problem. If the interval in question is the standard interval $\mathbf{I}$ and if $A = B = \{-1, 0, +1\}$ then for any continuous $f : \mathbf{I} \to \mathbf{I}$ is of the form $\tilde{m}$ for some (perhaps many) $m : A \to B$.

In our setting, the target is $\mathbb{W}$, hence the 1[st] condition above is irrelevant. If we specialize to the case that $A = \{-1, 0, +1\}$ and, as above, paths in $\{-1, 0, +1\}^{\mathbb{N}}$ name elements in $\mathbf{I}$ via symmetric ternary expansions (not—it must be emphasized—signed binary expansions) then given any $m : \{-1, 0, +1\}^\star \to \{-, +\}^\star$ covariant with respect to prolongation we obtain $\tilde{m} : \{-1, 0, +1\}^{\mathbb{N}} \to \mathbb{W}$. To obtain maps from $\mathbf{I}$ to $\mathbb{W}$ we still need the 2[nd] condition: two infinite paths naming the same element in $\mathbf{I}$ must be sent to the same element in $\mathbb{W}$ by $\tilde{m}$. Then continuity is automatic. For a proof, let w be a finite path in $\mathbb{W}$ and $x \in \mathbf{I}$ a point sent by $\tilde{m}$ to a prolongation of w. We seek a neighborhood of x all of which is sent by $\tilde{m}$ to prolongations of w. If x is named by a unique path in $A^\star$ let k be such that its initial word of length k is sent by m to a prolongation of w. Then, of course, the closed interval of all prolongations of that initial k-word are still sent to prolongations of w and x is an interior point of that closed interval. If, perchance, x is named by two infinite paths, we need only rephrase the argument. Let k by such that the initial word of length k of any infinite path that name x is sent to prolongations of w. The interval of all prolongations of those initial k-words are still sent to prolongations of w and x is an interior point of that closed interval. (We obtain a closed interval for each of the two names and x is an endpoint of each, the top end of one and the bottom end of the other, hence an interior point of their union.)

Finally, to show that it is an open continuous map, it suffices to find an open basis for $\mathbf{I}$ each member of which is mapped onto an updeal in $\mathbb{W}$. First, for each $w \in \{-1, 0, +1\}^\star$ note that the closed interval with $w, -1, -1, \ldots$ as lower bound and $w, +1, +1, \ldots$ as upper has a basic open set in $\mathbb{W}$ as image. But note that that remains true for the open interval with those same endpoints ($w, +1, +1, \ldots$ and $w, 0, 0, +1, +1, +1, \ldots$ are mapped to the same element in $\mathbb{W}$.) So, whenever an endpoint is other than $\pm\frac{1}{2}$ we delete it from the interval. The family of all such intervals is a basis.

The 1999 definition above of Wilson space was the most efficient way of communicating its total ordering (which ordering turns out to be implicit in its final coalgebra definition).[159]

The total ordering on $\mathbb{W}$ has endpoints and a canonical *total* interval coalgebra structure: $\top$-zooming sends every path, finite or infinite, with + as head to the rest of the word (its "tail") and sends all other words to the infinite path that's all minuses. $\bot$-zooming is, of course, defined dually. The induced map $\mathbb{W} \to \mathbf{I}$ may be described as the function that reads each path, finite or infinite, as a signed-binary expansion. Note that every element of $\mathbf{I}$ has a signed-binary expansion with no 0s, hence $\mathbb{W} \to \mathbf{I}$ is onto. Every element of $\mathbf{I}$ not an interior dyadic rational has a unique such expansion and every interior dyadic rational has exactly three expansions coming from $\mathbb{W}$.

We thus obtain an alternate characterization of the total ordering starting with a closed interval viewed as an ordered set and replacing each element of a countable dense subset of interior points with a three-element totally ordered set.

$\mathbf{I}$ has a canonical *partial*-coalgebra structure: partial $\top$-zooming is obtained by cutting the total $\top$-zooming operation down to a partial map, to wit, the one that's defined only on elements strictly larger than $\odot$. We treat, of course, $\bot$-zooming in the dual fashion. The induced map $\mathbf{I} \to \mathbb{W}$ followed by the previously induced map $\mathbb{W} \to \mathbf{I}$ is the identity map. The composition of the two induced maps in the other order is most easily described in terms of the last paragraph: it is the endomorphism on $\mathbb{W}$ obtained by collapsing to a point each of the inserted three-element sets (collapse every pair of elements with nothing between them).

36. *Addendum:* Vector Fields [2009–06–21]

Theorem 10.5 on the existence of standard models (p28) tells us that many things can be reduced to purely equational logic. For each natural number n consider the following Lipschitz extension of the theory of harmonic scales: n n-ary operators with equations

$$\bigvee_i (f_i\langle x_1, x_2, \ldots, x_n\rangle)^2 = \bigvee_i x_i^2$$

For each $i = 1, 2, \ldots, n$:

$$f_i\langle t^2x_1, t^2x_2, \ldots, t^2x_n\rangle = t^2 f_i\langle x_1, x_2, \ldots, x_n\rangle$$

[159] Note that the 1999 ordering is not the standard lexicographic ordering (which would take a finite word as prior to any of its prolongations) I said that there are two ordering of interest; the standard lexicographic ordering is not one of them.

And

$$\sum_i x_i f_i\langle x_1, x_2, \ldots, x_n\rangle = 0$$

If we view the fs as describing a self-map on n-space the first equation says that they describe a self-map that preserves the ℓ_∞-norm

The next n equations say that fs describe a self-map that's an ℓ^2-isometry on each ray.

The last equation says that the fs describe a self-map whose values are always orthogonal —in the standard ℓ_2 sense—to its arguments. (The summation is easily avoidable: pad out the two n-vectors with 0s to obtain vectors of dimension a power of two and replace adding with a binary tree of midpointing.)

Left out are equations to make the fs Lipschitz. Add them at will (with Lipschitz constant at least one.)

These equations are consistent iff n is even.

When n is even we have an easy model: define

$$f\langle x_1, x_2, \ldots, x_n\rangle = \langle x_2, -x_1, x_4, -x_3, \ldots, x_n, -x_{n-1}\rangle$$

For any n we may use a model of these equations to define a vector field on the $(n-1)$-sphere: for vectors of unit ℓ_2-norm simply define

$$g\langle x_1, x_2, \ldots, x_n\rangle = \frac{f\langle x_1, x_2, \ldots, x_n\rangle}{\|f\langle x_1, x_2, \ldots, x_n\rangle\|_2}$$

Because spheres of even dimension have no non-vanishing vector fields we know that these equations are inconsistent for any odd n. It is a sobering thought that beginning with the term $\top$ and using only a sequence of substitutions, each of which uses one of the defining equations with odd n, we can ultimately reach $\bot$. It is even more sobering to realize that Adams's theorem on parallelizable spheres [160]—which used just about all of algebraic topology then known [161]—was equivalent to a theorem about the consistency of certain finite families of equations.[162]

[160] "Vector fields on spheres," *Bull. Amer. Math. Soc.* 68 1962.

[161] In his first publication on the subject (*ibid.*) the parallelizability result appears as a special case of a (slightly later) more general result that used more than all of algebraic topology then known. In particular it used the Adams operations the construction of which required, conceptually, a quantification over functors; it was the first clearly important construction that did so. (Mac Lane never agreed with me that this was the most important single event in the history of category theory following its creation, indeed he omitted the Adams theorem from all of his histories of category theory: any actual application of category theory was somehow not to be considered category theory).

[162] Anyone can easily verify that in each of the three columns below the vectors are pairwise orthogonal.

$\langle +x_0 + x_1\rangle$	$\langle +x_0 + x_1 + x_2 + x_3\rangle$	$\langle +x_0 + x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7\rangle$
$\langle +x_1 - x_0\rangle$	$\langle +x_1 - x_0 + x_3 - x_2\rangle$	$\langle +x_1 - x_0 + x_3 - x_2 + x_5 - x_4 - x_7 + x_6\rangle$
	$\langle +x_2 - x_3 - x_0 + x_1\rangle$	$\langle +x_2 - x_3 - x_0 + x_1 + x_6 + x_7 - x_4 - x_5\rangle$
	$\langle +x_3 + x_2 - x_1 - x_0\rangle$	$\langle +x_3 + x_2 - x_1 - x_0 + x_7 - x_6 + x_5 - x_4\rangle$
		$\langle +x_4 - x_5 - x_6 - x_7 - x_0 + x_1 + x_2 + x_3\rangle$
		$\langle +x_5 + x_4 - x_7 + x_6 - x_1 - x_0 - x_3 + x_2\rangle$
		$\langle +x_6 + x_7 + x_4 - x_5 - x_2 + x_3 - x_0 - x_1\rangle$
		$\langle +x_7 - x_6 + x_5 + x_4 - x_3 - x_2 + x_1 - x_0\rangle$

It is beguilingly easy to believe that the middle column is a natural generalization of the left, and the right column of the middle. (If we erase everything except the indices we would be staring at tables for "nim-sum." The problem is just a matter of adjusting the signatures.) And clearly we should be able to obtain such a column of length 16. Adams became history's most conspicuous Fields non-Medalist by showing that if the top vector is $\langle x_0, x_1, \ldots, x_{n-1}\rangle$ with $n-1$ others given by continuous non-vanishing functions of the top and if the n vectors are pairwise orthogonal then $n = 1, 2, 4$ or 8.

37. *Addendum:* Extending the Reach of the Theorem on Standard Models

[2009–07–23]

We wish to obtain existence proofs for structures on spaces more general than standard n-cubes. We can describe any closed subset, C, of an n-cube, of course, as the kernel of a Lipschitz function, $M : \mathtt{I}^n \to \mathtt{I}$, defined on the n-cube (*e.g.*, $\sup_{y \in C} x \multimap\!\!\circ\, y$).

Suppose now that we wish to prove the existence of a fixed-point–free Lipschitz function $f : C \to C$. Given any Lipschitz f, Kirszbraun's theorem delivers a Lipschitz function $\tilde{f} : \mathtt{I}^n \to \mathtt{I}^n$ such that $\tilde{f}(x) = f(x)$ whenever it can, that is, whenever $x \in C$. [163] Thus we may formalize the notion of a self-map on C with an operator $f : \mathtt{I}^n \to \mathtt{I}^n$ satisfying the Horn sentence $\quad Mx = \top \;\Rightarrow\; M(fx) = \top$.

Because C is compact we know that if f is fixed-point–free then there's a $q \in \mathbb{I}_\star$ such that $\quad Mx = \top \;\Rightarrow\; fx \multimap\!\!\circ\, x \leq q$ (recall that $\mathbb{I}_\star$ are the elements below $\top$). But even if we prove the consistency of such a condition the theorem on standard models does not deliver a fixed-point–free map on C. The problem is the Horn condition: it's not an equation.[164]

Fortunately the Lipschitz condition on M allows us to find an equation equivalent to the Horn sentence:

$$Mx \leq (\top|)^{m+p}\left(M(\tilde{f}x)\right)$$

where m and p are integers that stipulate the Lipschitz condition on M and f:

$$(x \multimap\!\!\circ\, y) \leq \big((\top|)^m(Mx \multimap\!\!\circ\, My)\big) \;\text{ and }\; (x \multimap\!\!\circ\, y) \leq \big((\top|)^p(fx \multimap\!\!\circ\, fy)\big)$$

(Recall that any semiquation is equivalent to an equation).

It is clear that $M(x) \leq (\top|)^k(M\,(fx))$ implies the Horn sentence for any integer k. For the necessity, given $x \in \mathtt{I}^n$ let $x' \in C$ be such that $x \multimap\!\!\circ\, x'$ is maximal. Then $x \multimap\!\!\circ\, x' \leq (\top|)^p\left(\tilde{f}x' \multimap\!\!\circ\, \tilde{f}x\right) \leq (\top|)^{m+p}\left(M(\tilde{f}x') \multimap\!\!\circ\, M(\tilde{f}x)\right)$.

Thus if we replace the Horn sentence with this equational condition the consistency remains the same as does the existence of a standard model (Theorem 10.5, p28). (Alternatively, we could add an n-tuple of constants for each element of C and replace the Horn condition with the family of all conditions of the form $M(f\langle a_1, a_2, \ldots, a_n\rangle) = \top$.) [165]

The slogan: *consistent equational conditions have solutions.*

38. *Addendum:* Boolean-Algebra Scales [2009–10–30]

One of the better-known constructions in mathematics is that for group algebras: given a ring R and a group G, we construct $R[G]$ as the ring generated by R and the elements of G

[163] If one uses the standard Euclidean metric on $\mathtt{I}^n$ then Kirszbraun delivers an $\tilde{f}$ with the same Lipschitz constant as f. Given Hilbert spaces H_1, H_2 and a closed subset $C \subseteq H_1$ Kirszbraun extends any nonexpansive map $f : C \to H_2$ to all of H_1. Such easily generalizes from nonexpansive to Lipschitz but we're not done, we need to land in the standard cube. We do so by composing with the clearly nonexpansive map that sends each point in H_2 to its nearest point in the cube. (Such works for all convex closed sets, hence they all have nonexpansive retractions. It's worth noting the converse: the set of fixed-points for any nonexpansive map is necessarily convex.) Finally, since any norm on a finite dimensional real vector space is Lipschitz with respect to any other we obtain the advertised $\tilde{f}$.

[164] The theorem on standard models uses a simple quotient of the free algebra of an equational theory. Equations true for an algebra are true for quotient algebras. Not so for Horn sentences. Consider how easily the familiar Horn condition on rings $[x^2 = 0] \Rightarrow [x = 0]$ (*i.e.*, 'no nilpotents") fails to be preserved when passing to a quotient ring.

[165] There can be better ways for some special cases. For one example, we can effectively convert the n-cube into a (flat) n-torus by adding m equations for each new m-ary operator f, to wit, one for each $1 \leq i \leq m$:

$$f\langle x_1, \ldots, x_{i-1}, \bot|y\;, x_{i+1}, \ldots, x_m\rangle \;=\; f\langle x_1, \ldots, x_{i-1}, \top|y\;, x_{i+1}, \ldots, x_m\rangle.$$

subject to the relations that the multiplication induced on G by the ring structure coincides with that given by the group structure on G. Given a (chromatic, harmonic) scale S or any Lipschitz extension thereof and a Boolean algebra B we construct the **Boolean-algebra (chromatic, harmonic) scale $S[B]$** as the (chromatic, harmonic) scale generated by S and the elements of B subject to the relations that the lattice structure induced by the scale structure on the elements of B is as given by the Boolean-algebra structure (including, of course, that the top and bottom of B are the top and bottom of $S[B]$). Note that if x and y are a complementary pair in B they remain so in $S[B]$, hence all elements in B are extreme points in $S[B]$ and complementation in B coincides with dotting in $S[B]$.

The maps from $S[B]$ to any scale T are in natural correspondence with the pairs of maps, one a scale-homomorphism $S \to T$, the other a Boolean-homomorphism $B \to \mathcal{B}(T)$.[166] For the special case $S = \mathbb{I}$ we have that the functor $\mathbb{I}[-]$ from Boolean algebras to scales is the left adjoint of $\mathcal{B}(-)$ from scales to Boolean algebras. (More generally, if S is connected then $S[-]$ is the left adjoint of $\mathcal{B}(-)$ from the category of extensions of S to the category of Boolean algebras.)

When S is a connected scale the adjunction map from B to $\mathcal{B}(S[B])$ is an isomorphism. For a quick and dirty proof note that we may construct $S[B]$ as the S-valued continuous functions from the space of ultrafilters on B (where the topology on S is discrete). More generally (and constructively) we gain a handle on $S[B]$ as follows: we will say that a term is "pre-canonical" if it is of the form $(s_1 \wedge e_1) \vee (s_2 \wedge e_2) \vee \cdots \vee (s_n \wedge e_n)$ where $s_1, s_2, \ldots, s_n \in S$ and $\{e_1, e_2, \ldots, e_n\}$ is a partition of unity in B (pairwise disjoint and $e_1 \vee e_2 \vee \cdots \vee e_n = \top$); and "canonical" if, further, none of the e_is equals $\bot$ and the s_is are distinct. We need show only that every element of $S[B]$ is described by a canonical term, unique up to the ordering of the partition of unity.[167] For the existence of the term it suffices to show that elements named by canonical terms are closed under the scale operations: clearly any $s \in S$ is named by the canonical term $s \wedge \top$ and any $e \in B$ not in S (that is, other than $\top$ or $\bot$) by the canonical term $(\top \wedge e) \vee (\bot \wedge \dot{e})$; hence if the set of terms named by canonical terms are closed under the operations, then that set is necessarily all of $S[B]$. Zooming and—in the case of chromatic scales—the support operation are easy since each distributes with the lattice operations and fixes the extreme points: $((s_1 \wedge e_1) \vee (s_2 \wedge e_2) \vee \cdots \vee (s_n \wedge e_n))^\wedge = (\overset{\wedge}{s_1} \wedge e_1) \vee (\overset{\wedge}{s_2} \wedge e_2) \vee \cdots \vee (\overset{\wedge}{s_n} \wedge e_n)$ and, in the case of chromatic scales, $\overline{(s_1 \wedge e_1) \vee (s_2 \wedge e_2) \vee \cdots \vee (s_n \wedge e_n)} = (\overline{s}_1 \wedge e_1) \vee (\overline{s}_2 \wedge e_2) \vee \cdots \vee (\overline{s}_n \wedge e_n)$. The latter terms may be only pre-canonical but it is clear that every pre-canonical term is equal to a canonical term. Dotting requires a little work: we need a lemma on scales, to wit, that if the es form a partition of unity then $((s_1 \wedge e_1) \vee (s_2 \wedge e_2) \vee \cdots \vee (s_n \wedge e_n))^{\bullet} = (\dot{s}_1 \wedge e_1) \vee (\dot{s}_2 \wedge e_2) \vee \cdots \vee (\dot{s}_n \wedge e_n)$. The linear representation theorem for (chromatic, harmonic) scales comes to the rescue. Since the lemma is given by a family of universally quantified Horn sentences it suffices to check the equation on linear scales. But we know that in a linear scale all but one of the e_is will be $\bot$ and the one that is not will be $\top$ and that is quite enough. For midpointing suppose that $(s'_1 \wedge e'_1) \vee (s'_2 \wedge e'_2) \vee \cdots \vee (s'_m \wedge e'_m)$ is another canonical term. Then:

$$\bigvee_{i=1}^{n} (s_i \wedge e_i) \;\Big|\; \bigvee_{j=1}^{m} (s'_j \wedge e'_j) \;=\; \bigvee_{i=1}^{n} \bigvee_{j=1}^{m} \left((s_i | s'_j) \wedge e''_{i,j}\right)$$

where $e''_{i,j} = e_i \wedge e'_j$ (the "joint refinement" of the two given partitions of unity). Again, the easiest proof is simply to consider the equality in the case of a linear scale. The right-hand

[166] $\mathcal{B}$ was defined on p42.

[167] Note that if e_i were $\bot$ uniqueness would be lost: any s_i would do.

term is pre-canonical and—as before—can quickly be transformed into a canonical term. Given any Lipschitz extension this proof may be easily adapted, in particular for harmonic scales add the equation obtained by replacing midpointing with multiplication in the previous equation.

For the uniqueness of the canonical term note first that in a pre-canonical term for $\top$ it must be the case that $s_i = \top$ whenever $e_i \neq \bot$ (just specialize—once again—to any linear quotient). Hence if $(s_1 \wedge e_1) \vee (s_2 \wedge e_2) \vee \cdots \vee (s_n \wedge e_n) = (s'_1 \wedge e'_1) \vee (s'_2 \wedge e'_2) \vee \cdots \vee (s'_m \wedge e'_m)$ then we know that $s_i \multimapboth s'_j = \top$ whenever $e''_{i,j} \neq \bot$ which, in turn, says that $s_i = s'_j$ whenever $e_i \wedge e'_j \neq \bot$. We may thus infer that for every $1 \leqq i \leqq n$ there is $1 \leqq j \leqq m$ such that $s_i = s'_j$ and vice versa. Because the s_is in a canonical term are distinct this forces $n = m$ and the existence of a permutation π such that $s_i = s'_{\pi(i)}$. Since $e_i \wedge e'_j = \bot$ whenever $j \neq \pi(i)$ we may infer that $e_i = e'_{\pi(i)}$.[168]

Scales of the form $\mathtt{I}[B]$ thus succeed in mixing Aristotle, Łukasiewicz and Girard: for extreme points $\overset{\vee}{|} = ⅋ = \vee$ and $\overset{\wedge}{|} = \otimes = \wedge$ hence $x \multimap y$ is the same as the standard material implication; the support and co-support operators translate from Łukasiewicz and Girard to Aristotle ($\overline{\hat{x}}$ is the assertion, for example, that x is more probable than not).

39. *Addendum:* On the Definition of I-Scales [2010–07–28]

Our definition of I-scales should be considered as the minimal definition needed to make it both unique on the standard interval and co-congruent with the theory of scales. It has a unique maximal equational consistent extension.

In Section 3 (p12–13) we used the fact that a particular scale ($\mathbb{I}$) appears as a subscale of every non-trivial minor scale to imply the existence of a unique maximal equational consistent extension. In this case we use the fact that a particular scale ($\mathtt{I}$) appears as a *quotient* scale of every non-trivial I-scale: if any equation is consistent with the theory of I-scales then it has a non-trivial model which, in turn, has a quotient scale isomorphic (as an I-scale) to I, hence holds for I. Among such equations are:

$$\odot x = \odot \qquad (r|s)x = rx|sx \qquad \top x = x \qquad (rs)x = r(sx)$$ [169]

It was never clear to the writer just how I-scales should be defined. The minimal definition was the choice. But note that the every theorem herein holds equally well for the maximal definition.

40. *Addendum:* Proofs for Section 0 [2011–06–02]

At the beginning of Section 4 (p14–17) on lattice structure we explained how a closed-interval homomorphism from $\mathcal{C}(X)$ to I necessarily preserves order. Clearly it also sends functions constantly equal to an element in $\mathbb{I}$ to that same element in $\mathbb{I}$. Combine with the covariance and we have that it sends functions constantly equal to any element in I to that same element in I. From that we may conclude that if $f : X \to \mathtt{I}$ is bounded above/below by $x \in \mathtt{I}$ than it is sent to an element in I also bounded above/below by x.[170]

[168] Having found this construction for scales, I assume that it is ancient knowledge that the analog construction for Boolean algebras of central idempotents works as well. The case $\mathbb{Z}_2[B]$ returns us, of course, to Boole's original *Laws of Thought.*

[169] Note that in the left-hand terms $\odot, r|s, \top, rs$ refer to the structure of I and in the right-hand terms $\odot$ and $rx|sx$ refer to the structure in the given I-scale. (The right-hand term $r(sx)$ uses composition of unary operations.)

[170] We will not need it here, but if one now adds the fact that any closed-interval homomorphism necessarily preserves dotting (since $\dot{u}$ is uniquely characterized by $\dot{u}\,|u = \bot|\top$) to obtain that if $\mathcal{C}(X)$ and $\mathcal{C}(Y)$ are viewed as metric spaces under the uniform norm then any closed-interval map between them is necessarily nonexpansive.

For $f \in \mathcal{C}(X)$ define its **variance**, by $\mathbf{var}(f) = \sup(f) - \inf(f)$. It follows that any closed-interval endomorphism on $\mathcal{C}(X)$ conserves variance, that is, does not increase it. A special case, note, is that a convex combination has variance at most equal to the greatest of the variances of the functions of which it is a combination.

In the case $X = \mathbf{I}$ we may rewrite the 4$^{\text{th}}$ equation [171] (using the 3$^{\text{rd}}$ equation) as:

$$⨍ f(x)\,\mathrm{d}x \quad = \quad ⨍ f(\bot|x) \mid f(\top|x)\ \mathrm{d}x$$

Let $T : \mathcal{C}(\mathbf{I}) \to \mathcal{C}(\mathbf{I})$ denote the operator such that $(Tf)(x) = f(\bot|x) \mid f(\top|x)$. Note first, that T is a closed-interval endomorphism.[172] It suffices to show that for any $f \in \mathcal{C}(\mathbf{I})$ the sequence $\{\inf(T^n f)\}$ is non-decreasing and $\{\sup(T^n f)\}$ is non-increasing and, finally, that their difference converges to 0. We've already noted that T as any convex combinator is monotonic on both inf and sup. Hence it suffices to show that $\lim \mathrm{var}(T^n f) = 0$.

We need a description of $T^n f$. It is the arithmetic mean of 2^n functions each of which is of the form

$$\mathbf{I} \xrightarrow{e_1|} \mathbf{I} \xrightarrow{e_2|} \mathbf{I} \to \cdots \mathbf{I} \xrightarrow{e_n|} \mathbf{I} \xrightarrow{f} \mathbf{I}$$

where each e_i is either $\bot$ or $\top$. Each of the 2^n functions is a contraction down to a subinterval 2^{-n} of the length of $\mathbf{I}$ followed by f. It thus suffices to show that for any $\varepsilon > 0$ we can find n such that when f is restricted to any of those subintervals its variance is bounded by ε. Suppose there were no such n. Consider the binary tree of intervals (partially ordered by containment) arising from all the T^ns; throw away all those intervals on which f is of variance bounded by ε; if the resulting tree were infinite then—as for any infinite rooted tree with finite branching—König's lemma would give us an infinite path, that is, an infinite chain of closed intervals—each half the length of the previous—on which f is of variance greater than ε. But f could not be continuous at the point that lies in their intersection. [173]

This proof easily generalizes to $\mathbf{I}^m$. In the case $m = 2$ replace the 4$^{\text{th}}$ equation with two equations:

$$⨍ f\langle\bot|x, y\rangle\,\mathrm{d}x \mid ⨍ f\langle\top|x, y\rangle\,\mathrm{d}x \quad = \quad ⨍ f\langle x, y\rangle\,\mathrm{d}x \quad = \quad ⨍ f\langle x, \bot|y\rangle\,\mathrm{d}x \mid ⨍ f\langle x, \top|y\rangle\,\mathrm{d}x$$

[171] The 4$^{\text{th}}$ equation re-occurred to me via an outrageously circuitous route (I say "re-occurred" because I had quite forgotten the diagram that illustrated it in my 1999 lecture mentioned in Section 32, p75–77). We could, of course, convert integration on $\mathbf{I}$ to integration on the circle: given $f : \mathbf{I} \to \mathbf{I}$ we could throw away one of the endpoints and view the resulting half-open interval as a circle torn apart. The problem is that we will have converted the integration of a continuous function on $\mathbf{I}$ to the integration of a function on the circle with a jump-discontinuity. But clearly the function whose value at x is $f(\dot{x})$ should have the same mean-value, hence instead of integrating $f(x)$ we could integrate $f(\dot{x})|f(x)$; its conversion to a circle does yield a continuous function. But then we note that we're integrating a function symmetric on $\mathbf{I}$ and it would suffice to find the mean-value on either half. Choose the top half. What we've done is replace the mean value of $f(x)$ with the midpoint of the mean values of $f(\top|x)$ and $f\left((\top|x)^{\cdot}\right)$. At a lecture at the University of Cambridge I started calling this the "puff-pastry" method: roll out the dough; coat with butter; fold over; repeat until done. Perhaps it was because I was so happy with that name that it took me so long to note that $(\top|x)^{\cdot}$ could be replaced with $(\bot|\,\dot{x})$ and that the mean value of $f(\bot|\,\dot{x})$ would be the same without the dot.

[172] What we are showing, therefore, is that the mean-value operator from $\mathcal{C}(\mathbf{I})$ to $\mathbf{I}$ is characterized as the closed-interval homomorphism that is constant on the orbits of T.

[173] We used a contradiction to show that a particular procedure works (to wit, trying each n in sequence). What will be viewed by some as a more constructive proof for the existence of such an n start by taking $\mathbf{I}$ to be the standard interval, define $\mathrm{var}_{[x,y]}(f)$ to be the variance of f on the closed interval with endpoints x and y, and let $h(x) = \sup\{\, y : \mathrm{var}_{[x,y]}(f) \le \varepsilon/2 \,\}$. Denote $h^i\bot$ as a_i, let $a_\infty = \sup_i a_i$ and (using the continuity of f at a_∞) let $m = \min\{\, i : \mathrm{var}\langle a_i,\, a_\infty\rangle \le \varepsilon/2 \,\}$. Then necessarily $a_{m+1} \ge a_\infty = \top$ (the unique fixed-point of h). There's a unique dyadic rational of minimal denominator in the open interval $(a_i,\, a_{i+1})$ for each $i \le m$. Let 2^n be the largest of those minimal denominators. The number being sought is n. (For any pair of subsets A, B with $A \cap B \neq \emptyset$ note that $\mathrm{var}_{A\cup B}(f) \le \mathrm{var}_A(f) + \mathrm{var}_B(f)$ any f, any domain.)

For each m the 4^{th} equation is replaced by m equations, one for each dimension: for each $i = 1, 2, \ldots, m$ there will be integrands $f\langle x_1, \ldots, \bot | x_i, \ldots, x_m\rangle$ and $f\langle x_1, \ldots, \top | x_i, \ldots, x_m\rangle$. The operator T becomes a Convex of 2^m functions (and for T^n it's 2^{mn}) and the binary tree becomes a tree in which each vertex has 2^m branches. [174]

For the case that X is a a compact group, G, we need a few definitions. Given $f \in \mathcal{C}(G)$ let R_f be the set of all convex combinations of the right-translates of f [175] and $\overline{R}_f$ the closure of R_f (using, of course, the uniform norm on $\mathcal{C}(G)$). It seems to this writer that one needs two great von-Neumann ideas. The first is that $\overline{R}_f$ contains a constant function. A similar argument yields that the closure, $\overline{L}_f$, of L_f, the set of convex combinations of left-translates of f also has a constant function. We will not need a great von-Neumann idea to establish (below) that any constant function in $\overline{R}_f$ is equal to any constant function in $\overline{L}_f$ which more than suffices to establish the uniqueness of both.[176] The fact that the constant values of those unique constant functions yields a closed-interval homomorphism is also easy: given $f_1, f_2 \in \mathcal{C}(G)$ we will let $T_1, T_2 : \mathcal{C}(G) \to \mathcal{C}(G)$ each denote an operator that delivers a given convex combination (to be chosen) of a given sequence of right-translates (to be chosen) of its argument; choose T_1 so that $T_1 f_1$ is within ε of the unique constant function in $\overline{R}_{f_1}$ and choose T_2 so that $T_2(T_1 f_2)$ is within ε of the unique constant function in $\overline{R}_{T_1 f_2} \subseteq \overline{R}_{f_2}$; it is still the case that $T_2(T_1 f_1)$ is within ε of the unique constant function in R_{f_1}; thus $T_2(T_1(f_1|f_2)) = (T_2(T_1(f_1)) \mid (T_2(T_1(f_2))$ is within ε of the mean of constant functions in $\overline{R}_{f_1}$ and $\overline{R}_{f_2}$.

To find a constant function in $\overline{R}_f$ we need the second great von-Neumann idea, to wit, use that $\overline{R}_f$ is compact. We can then find $h \in \overline{R}_f$ of minimal variance and will finish by showing that $\mathrm{var}(h) > 0$ would allow an easy construction of an element of smaller variance.

So we start with the compactness argument. For any continuous $f : G \to \mathbf{I}$ the induced function $G \to \mathcal{C}(G)$ that sends $\alpha \in G$ to f^α is continuous (using the uniform norm in G). The proof requires a little work (it will use, for example, the compactness of G [177]). Given $\varepsilon > 0$ we need an open neighborhood U of $1 \in G$ such that $\|f - f^\alpha\| < \varepsilon$ all $\alpha \in U$. If there were no such neighborhood then Moore-Smith convergence works well for reaching a contradiction: from the (down-)directed set of neighborhoods of 1 we can construct a net by sending U to $y \in G$ such that $|f(y) - f^\alpha(y)| \geqq \varepsilon$ for some $\alpha \in U$; the compactness of G allows us to find a sub-net such that the ys converge to a point $z \in G$ such that for every U there's an $\alpha \in U$ with $|f(z) - f(z\alpha)| \geqq \varepsilon$ which belies the continuity of f.

Thus the set of right-translates of f, being the image of a continuous map from a compact space, is itself a compact subset of $\mathcal{C}(G)$. For any $\varepsilon > 0$ we may choose a finite sequence $\{f^{\alpha_1}, f^{\alpha_2}, \ldots, f^{\alpha_n}\}$ which is $\varepsilon/2$-dense in the set of all right-translates (that is, every right-translate is within distance bounded by $\varepsilon/2$ from one of the f^αs).

The set of convex combinations of the f^αs is $\frac{\varepsilon}{2}$-dense in R_f by the following argument: given $r_1 f^{\beta_1} + r_2 f^{\beta_2} + \cdots + r_m f^{\beta_m}$ replace each of the f^βs with one of the $\frac{\varepsilon}{2}$-close f^αs to obtain

[174] Without mentioning mean-values we could have stated that for any coninuous $f : \mathbf{I} \to \mathbf{I}$ the sequence of functions obtained by iterating T uniformly converges to a constant funcion. If the domain is a cube of dimension m and if T_j defines the operator that applies the T-operator to the jth coordinate then any sequence of T_js in which each T_j appears infinitely often for each j from 1 through m likewise will yield a sequence of functions uniformly converging to a constant function on $\mathbf{I}^m$.

[175] For our setting we need convex combinations using dyadic rationals, but such barely affects the proof. Indeed, one may simply understand the phrase "convex combination" in this section to mean one that uses only dyadic rationals.

[176] More than suffices because, of course, it also establishes that right-invariance of integration on compact groups is equivalent to left-invariance.

[177] If we take $\mathbf{I}$ to be the standard interval, G to be the group of positive reals under multiplication and $f \in \mathcal{C}(G)$ to be defined by $f(x) = \cos(\pi x)$ then the sequence $\{\alpha_n = 1 + 1/n\}_{n=1}^{\infty}$ converges to 1 in G but $\|f - f^{\alpha_n}\|_\infty$ is constantly equal to 2 (because $(-1)^n = f(n) = -f^{\alpha_n}(n)$). One may make this example look more complicated by using, instead, the (isomorphic) group of reals under addition. But the more complicated version does have its merits: it's easy to see that the distances between adjacent critical points of $\cos(\pi e^x)$ are arbitrarily small.

a convex combination of the f^αs (albeit, one with repeated f^αs). This convex combination is of distance bounded by $\frac{\varepsilon}{2}$ from the arbitrary one. We can then finish, of course, by reordering the summands and using the distributive law to reduce down to a sum of multiples of at most n of the f^αs.

But the set of all convex combinations (dyadic rational or not) of the f^αs is itself compact (it is the image of a map from the unit simplex in n-space, that is, the set of n-tuples of non-negative reals that add to 1.) hence has also a finite $\frac{\varepsilon}{2}$-dense subset which, perforce, is ε-dense in the set of all convex combinations of the right-translates of f. All of which establishes that R_f is totally bounded. In any complete metric space the closure of a totally bounded set is, of course, compact and $\overline{R}_f$ has an element of minimal variance.

We could now easily obtain the existence of a constant function in $\overline{R}_f$ by taking it as an element of minimal variance. It is even easier to take it as an element of minimal maximum value; we need only show that for any non-constant function h there is a convex combination of h-translates with smaller maximum. Let $U \subseteq G$ be the open set $\{\, x : h(x) < \max h \,\}$. Let $\alpha_1, \alpha_2, \ldots, \alpha_n$ be such that $G = U\alpha_1^{-1} \cup U\alpha_2^{-1} \cup \cdots \cup U\alpha_n^{-1}$. Then the maximum value of the convex combination $h^{\alpha_1}|(h^{\alpha_1}|(\cdots|(h^{\alpha_n})\cdots))$ is less than $\max h$ (it suffices to find for any $x \in G$ an i such that $h^{\alpha_i}(x) < \max h$ but $x\alpha_i \in U$ is—obviously—equivalent to $x \in U\alpha_i^{-1}$).

Finally the advertised easy proof of uniqueness. Suppose that $\overline{L}_f$ has a function constantly valued a and $\overline{R}_f$ one constantly valued $b \neq a$. Replace f with $\dot{f}$, if necessary, to guarantee that $a < b$. We know that there is convex combination of left-translates of f whose maximum value is strictly less than $a|b$ and a convex combination of right-translates whose minimum value is strictly greater than $a|b$. We introduce a symbolic arithmetic just for this argument. Given an arbitrary $g \in \mathcal{C}(G)$ and the qs and αs that are used for the convex combination in L_f with max less than $a|b$ we understand the formal product $(q_1^{\alpha_1} + q_2^{\alpha_2} + \cdots + q_m^{\alpha_m})g$ to be $q_1\,{}^{\alpha_1}g + q_2\,{}^{\alpha_2}g + \cdots + q_m\,{}^{\alpha_m}g$ where the ${}^{\alpha_i}g$s denotes left-translates.

Given the rs and βs that describe the convex combination in R_g with min greater than $a|b$ we understand the formal product $g({}^{\beta_1}r_1 + {}^{\beta_2}r_2 + \cdots + {}^{\beta_n}r_n)$ to be $g^{\beta_1}r_1 + g^{\beta_2}r_2 + \cdots + g^{\beta_n}r_n$. Then it does not matter how we associate the formal triple product

$$(q_1^{\alpha_1} + q_2^{\alpha_2} + \cdots + q_m^{\alpha_m})f({}^{\beta_1}r_1 + {}^{\beta_2}r_2 + \cdots + {}^{\beta_n}r_n)$$

In each case we obtain the sum

$$\sum_{i,j} q_i\,{}^{\alpha_i}f^{\beta_j}\,r_j$$

where, of course, ${}^{\alpha_i}f^{\beta_j}$ is the function that sends x to $f(\alpha_i x \beta_j)$. Since

$$\Big((q_1^{\alpha_1} + q_2^{\alpha_2} + \cdots + q_m^{\alpha_m})f\Big)({}^{\beta_1}r_1 + {}^{\beta_2}r_2 + \cdots + {}^{\beta_n}r_n)$$

is a convex combination of functions all of whose maximum values are strictly less than $a|b$ and

$$(q_1^{\alpha_1} + q_2^{\alpha_2} + \cdots + q_m^{\alpha_m})\Big(f({}^{\beta_1}r_1 + {}^{\beta_2}r_2 + \cdots + {}^{\beta_n}r_n)\Big)$$

is a convex combination of functions all of whose minimum values are strictly greater than $a|b$ we obtain the desired contradiction.

41. *Addendum:* Extreme Points, Faces and Convex Sets [2011–08–22]

Given a binary operation $\diamond$ on a set S, to say that a subset S' is closed under $\diamond$ means that $x, y \in S'$ implies $x \diamond y \in S'$. We'll say here that S' is **co-closed** if the converse holds: $x \diamond y \in S'$ implies $x, y \in S'$ [178] and **bi-closed** if both closed and co-closed.

We defined a subset of a scale to be a face if it bi-closed under midpointing. In the theory of convex sets the usual condition is that a subset F is a face if it is bi-closed under all convex binary combinations, that is, for all x, y in the convex set and $0 < a < 1$ it is the case that $ax + (1-a)y \in F$ iff $x, y \in F$.[179] A convex set is usually understood to be a subset of a real vector space and in that case we will show below that the two definitions of face are—fortunately—equivalent. The proof requires the Archimedean condition (if K is a non-Archimedean harmonic scale note that the Jacobson radical of any non-trivial closed interval is a $\top$-face but not co-closed under all convex binary operations).

No special conditions are required to show that bi-closure under midpointing implies convexity: given a convex combination $ax + by$ where x and y are in a midpoint-face F, then closure under midpointing says that $x|y \in F$ and co-closure under midpointing then says that $ax + by \in F$ because, easily enough, $(ax + by) \bigm| (bx + ay) = x|y$.[180] The problem is the co-closure.

Given positive a, b such that $a + b = 1$ we do have—even in the non-Archimedean case—that if $ax + by \in F$ then either $x \in F$ or $y \in F$, to wit, the one with the larger coefficient: since if $0 \leqq a \leqq b$ and $a + b = 1$ then $ax + by = \bigl(2ax + (1-2a)y\bigr) \,|y$ and midpoint co-closure yields that $y \in F$. But we have more: $2ax + (1-2a)y \in F$. By iteration we obtain that if $0 \leqq a \leqq b$ and $a + b = 1$ then it the case that $2a^n x + (1 - 2a^n)y \in F$ for all $n \geqq 0$ such that $2^n a \leqq 1$. In the Archimedean case we have a largest such n and hence a case where the coefficient of x is the larger coefficient and that yields $x \in F$.

In the one-element case—that is, in the case of an extreme point—we do not need the Archimedean condition. If $\{e\}$ is co-closed under midpointing and if $0 < a \leqq b$, $a + b = 1$ and $ax + by = e$ then we have just seen that $y = e$. But from $ax + (1-a)e = e$ we infer that $ax = ae$ hence that $x = e$.

42. *Addendum:* Scale Spectra Are Compact Normal [2013–04–02]

Near the end of the discussion of Theorem 8.6, in the TAC version, I wrote in reference to Spec(S):

> It is always a spatial locale: the points are the "prime" congruences, that is, those that are not the intersection of two larger congruences. Translated to filters: $\mathcal{F}$ is a point if it has the property that whenever $x \vee y \in \mathcal{F}$ it is the case that either $x \in \mathcal{F}$ or $y \in \mathcal{F}$. Put another way, of course, the points of Spec(S) are the linearly ordered quotients of S. We will show that Spec(S) is compact normal (but not always Hausdorff).

Alas, I forgot that last promise.

First, note that the definition of prime is not quite right. We need to exclude the entire filter. (One way to do so is to replace the word "two" with "a finite number of.")

[178] Note that the weaker condition, $x \diamond y \in S'$ implies either $x \in S'$ or $y \in S'$, holds iff the compliment of S' is closed in which case we could say that S' is **open** under $\diamond$.
[179] An easy inductive argument shows that the closure and co-closure of binary convex combinations each implies the same for n-ary convex combinations.
[180] We need both midpoint closure and co-closure to obtain convex closure. $\mathbb{I}$ is midpoint-closed but not, of course, convex-closed in $\mathbf{I}$.

Just as in the ancestral subject, the topology is given by a basis whose members are of the form

$$U_s = \left\{ \mathcal{F} \in \mathrm{Spec}(S) \, : \, s \notin \mathcal{F} \right\}$$

where $s \in S$. We obtain the same list of identities as in Section 23 (p51–53):

$$\begin{aligned} U_s \cap U_t &= U_{s \vee t} \\ U_s \cup U_t &= U_{s \wedge t} \\ U_s &= U_{\hat{s}} \\ U_\top &= \emptyset \\ U_\perp &= \mathrm{Spec}(S) \end{aligned}$$

The compactness argument for Spec(S) is also a repetition of that in Section 23, p51–53 (for Max(S)): given a subset $S' \subseteq S$ the necessary and sufficient condition that the family of sets $\{ U_s \, : \, s \in S' \}$ be a cover of Spec(S) is that every prime zoom-invariant filter excludes some element of S', or—put another way—that no prime zoom-invariant filter contains all of S'. Since every zoom-invariant filter is the intersection of the prime zoom-invariant filters that contain it, the condition is equivalent to the zoom-invariant filter generated by S' being entire, that is, the condition that $\perp$ be in the set obtained by closing S' under finite meet and zooming. But, of course, only a finite number of elements of S' can be involved in any such demonstration and thus their corresponding basic open sets yield the finite subcover.

Normality translates to the condition on the locale of open sets of a space X, that for any open sets V, W such that $V \cup W = X$ there exist open sets V', W' such that: [181]

$$\begin{aligned} V &\supseteq V' \\ W &\supseteq W' \\ V' \cap W' &= \emptyset \\ V' \cup W &= X \\ V \cup W' &= X \end{aligned}$$

Using compactness the general case reduces to the case where V and W are finite unions of basic open sets. [182] In the case at hand we have a basis closed under finite unions. Hence we may assume that there are elements $s, t \in S$ such that V and W are of the form U_s and U_t. And it is clearly sufficient (and—left as an exercise—necessary) to find V', W' of the form $U_{s'}, U_{t'}$. The condition $U_s \cup U_t = \mathrm{Spec}(S)$ is equivalent to $U_{s \wedge t} = \mathrm{Spec}(S)$ which is, in turn, equivalent to the condition that a finite iteration of $\top$-zooming pushes $s \wedge t$ down to $\perp$. Since zooming distributes with the lattice operations we may replace s, t with the result of that finite iteration. Hence we may assume that $s \wedge t = \perp$. The same maneuver applies to s', t'. The condition $V \supseteq V'$ translates, of course, to $V \cap V' = V'$, which after sufficient $\top$-zooming is equivalent to $s \vee s' = s'$, that is, $s \leq s'$. Hence the normality of Spec(S) would be a consequence of the condition that given $s, t \in S$ such that $s \wedge t = \perp$ there exist $s', t' \in S$ such that:

$$\begin{aligned} s &\leq s' \\ t &\leq t' \\ s' \vee t' &= \top \\ s' \wedge t &= \perp \\ s \wedge t' &= \perp \end{aligned}$$

[181] The standard definition, of course, is that any pair of disjoint closed sets has a pair of disjoint open neighborhoods. Take V and W as the complements of the closed sets, $V'(W')$ as the neighborhood of the complement of $W(V)$.

[182] The entire space is covered by the basic open sets that are contained either in V or in W. Chose a finite subcovering. Replace each of V and W with the union of those basic sets that are contained therein. The union of these two replacements is still X and any pair V', W' that satisfies the five conditions for these replacements automatically satisfies them for the originals.

Using (for the last time) that $U_s = U_t$ we replace the last two equalities with

$$\begin{array}{rcl} \widehat{s' \wedge t} & = & \bot \\ \widehat{s \wedge t'} & = & \bot \end{array}$$

We thus finish with the observation that $s' = t \multimap s$ and $t' = s \multimap t$ do just what is needed:

$$\begin{array}{ccccccccc} s & = & \bot \overset{\vee}{|} s & \leqq & \dot{t} \overset{\vee}{|} s & = & t \multimap s & = & s' \\ t & = & \bot \overset{\vee}{|} t & \leqq & \dot{s} \overset{\vee}{|} t & = & s \multimap t & = & t' \\ & & s' \vee t' & = & (t \multimap s) \vee (s \multimap t) & = & \top & & \\ \widehat{s' \wedge t} & \leqq & s' \overset{\wedge}{|} t & = & (t \multimap s) \overset{\wedge}{|} t & = & t \wedge s & = & \bot \\ \widehat{s \wedge t'} & \leqq & s \overset{\wedge}{|} t' & = & s \overset{\wedge}{|} (s \multimap t) & = & s \wedge t & = & \bot \end{array}$$

The middle row uses, of course, the equation of linearity and the last two rows the construction of the meet operation (and the fact that $x \wedge y = (x \wedge y)|(x \wedge y) \leqq x|y$).

43. *Addendum:* Signed-Binary Automata [2013–10–25]

Given terms "$t_1|(t_2|(t_3|(\cdots$" and "$u_1|(u_2|(u_3|(\cdots$" their midpoint is "$v_1|(v_2|(v_3|(\cdots$" where $v_i = t_i|u_i$. We can easily construct an instantaneous automaton that produces a stream of symbols from the set $\{\bot, \dot\smile, \odot, \dot\frown, \top\}$ where the **quarter-moons** are defined by $\dot\smile = \bot|\odot$ and $\dot\frown = \odot|\top$. The problem, then, is to remove those quarter-moons. Consider:

$$\begin{array}{rclcrcl} \dot\frown\big|(\top|x) & = & \top\big|(\odot|x) & \qquad & \dot\smile\big|(\top|x) & = & \odot\big|(\odot|x) \\ \dot\frown\big|(\dot\frown|x) & = & \top\big|(\dot\smile|x) & & \dot\smile\big|(\dot\frown|x) & = & \odot\big|(\dot\smile|x) \\ \dot\frown\big|(\odot|x) & = & \odot\big|(\top|x) & & \dot\smile\big|(\odot|x) & = & \odot\big|(\bot|x) \\ \dot\frown\big|(\dot\smile|x) & = & \odot\big|(\dot\frown|x) & & \dot\smile\big|(\dot\smile|x) & = & \bot\big|(\dot\frown|x) \\ \dot\frown\big|(\bot|x) & = & \odot\big|(\odot|x) & & \dot\smile\big|(\bot|x) & = & \bot\big|(\odot|x) \end{array}$$ [183]

The significance of these computations is that each term may be replaced with one whose left-most symbol is not a quarter-moon and that allows us to build an automaton that systematically removes all quarter-moons from the output streams. The remarkably simple finite automaton that results for midpointing signed-binary expansions has just three states. Its nine pairs of input digits behaviorally group themselves into five classes defined by the sum of the two digits, hence the inputs will be denoted with $-2, -1, 0, +1 + 2$. The states will be $\mathsf{S}_-, \mathsf{S}_\circ$ (the initial state) and S_+.

[183] An easy corollary of Theorem 15.1 (p34) but for the fastidious (see Section 46 (p114–119) for subscorings):

$$\dot\frown\big|(\top|x) = (\top|\odot)\big|(\top|x) = \top\big|(\odot|x)$$

$$\dot\frown\big|(\dot\frown|x) = (\odot|\top)\big|(\dot\frown|x) = (\odot|\dot\frown)\big|(\top|x) =$$
$$\Big((\top|\bot)\big|(\top|\odot)\Big)\big|(\top|x) = \Big(\top\big|(\bot|\odot)\Big)\big|(\top|x) = (\top|\dot\smile)\big|(\top|x) = \top\big|(\dot\smile|x)$$

$$\dot\frown\big|(\odot|x) = (\odot|\top)\big|(\odot|x) = \odot\big|(\top|x)$$

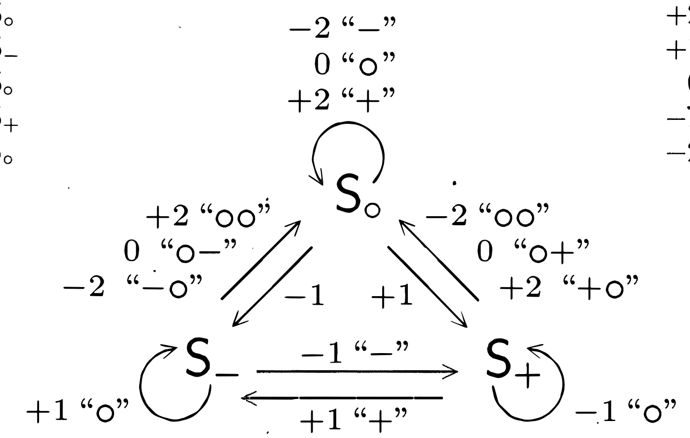
$$\dot\frown\big|(\dot\smile|x) = (\odot|\top)\big|(\dot\smile|x) = (\odot|\dot\smile)\big|(\top|x) =$$
$$\Big((\bot|\top)\big|(\bot|\odot)\Big)\big|(\top|x) = \Big(\bot\big|(\top|\odot)\Big)\big|(\top|x) = (\bot|\dot\frown)\big|(\top|x) = (\bot|\top)\big|(\dot\frown|x) = \odot\big|(\dot\frown|x)$$

$$\dot\frown\big|(\bot|x) = (\top|\odot)\big|(\bot|x) = (\top|\bot)\big|(\odot|x) = \odot\big|(\odot|x)$$

The **signed-binary midpoint automaton**:

Next State

	S_-	$S_\circ$	S_+
+2	$S_\circ$	$S_\circ$	$S_\circ$
+1	S_-	S_+	S_-
0	$S_\circ$	$S_\circ$	$S_\circ$
−1	S_+	S_-	S_+
−2	$S_\circ$	$S_\circ$	$S_\circ$



Output

	S_-	$S_\circ$	S_+
+2	oo	+	+o
+1	o		+
0	o−	o	o+
−1	−		o
−2	−o	−	oo

The states may be interpreted as follows: in state $S_\circ$ the midpoint of the input streams (so far) is equal to the present output. Whenever we leave $S_\circ$ a stammer occurs and the machine moves to either S_+ or S_-; in S_+ the midpoint of the present input streams is larger; in S_- it's smaller. Whenever we return to $S_\circ$ a stutter occurs: two output digits. The machine is never more than one output digit behind the number of input pairs (that is, between every pair of stammers there's a stutter).[184]

It is worth noting that if both input streams are without bad tails then so is the output: whenever returning to $S_\circ$ a "o" is produced hence an output bad tail would require the machine eventually to stay entirely in the two lower states, S_+ and S_-, or eventually to stay entirely in the single state $S_\circ$; in the 1st case the output would not have any adjacent pairs of +s or −s; in the 2nd case a bad tail would be produced only if both input streams become eventually all +s or eventually all −s (and if neither has a bad stream then both converge to $\top$ or both to $\bot$).

---

[184] A non-stuttering machine (with six states, one of which, I, is strict initial) is available:

Next State

	I	S_{-2}	S_{-1}	S_0	S_{+1}	S_{+2}
+2	S_{+2}	S_{+2}	S_0	S_{+2}	S_0	S_{+2}
+1	S_{+1}	S_{+1}	S_{-1}	S_{+1}	S_{-1}	S_{+1}
0	S_0	S_0	S_{-2}	S_0	S_{+2}	S_0
−1	S_{-1}	S_{-1}	S_{+1}	S_{-1}	S_{+1}	S_{-1}
−2	S_{-2}	S_{-2}	S_0	S_{-2}	S_0	S_{-2}

Output

	I	S_{-2}	S_{-1}	S_0	S_{+1}	S_{+2}
+2		−	o	o	+	+
+1		−	o	o	+	+
0		−	o	o	o	+
−1		−	−	o	o	+
−2		−	−	o	o	+

A stammer occurs at the very beginning, thereafter it is always exactly one output digit behind the number of input pairs; alternatively, a machine that doesn't stammer until it has to:

Next State

	I	S_{-2}	S_{-1}	S_0	S_{+1}	S_{+2}
+2	I	S_{+2}	S_0	S_{+2}	S_0	S_{+2}
+1	S_{+1}	S_{+1}	S_{-1}	S_{+1}	S_{-1}	S_{+1}
0	I	S_0	S_{-2}	S_0	S_{+2}	S_0
−1	S_{-1}	S_{-1}	S_{+1}	S_{-1}	S_{+1}	S_{-1}
−2	I	S_{-2}	S_0	S_{-2}	S_0	S_{-2}

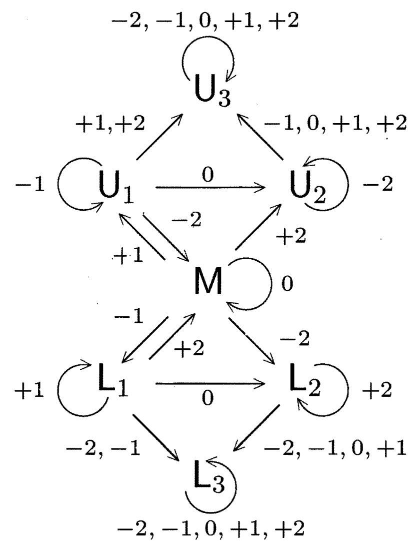
Output

	I	S_{-2}	S_{-1}	S_0	S_{+1}	S_{+2}
+2	+	−	o	o	+	+
+1		−	o	o	+	+
0	o	−	o	o	o	+
−1		−	−	o	o	+
−2	−	−	−	o	o	+

The lattice operations require, curiously, an automaton larger than the midpoint operation. First—for reasons to become clear—consider an automaton with seven states denoted $\mathsf{L}_3, \mathsf{L}_2, \mathsf{L}_1, \mathsf{M}$ (the initial state),$\mathsf{U}_1, \mathsf{U}_2, \mathsf{U}_3$. We imagine an "upper" stream of signed binary digits and a "lower" one. The nine input pairs again group themselves into five classes for the next-state behavior. It is determined by their difference: the upper digit minus the lower digit. The **signed-binary lattice automaton**;

Next State

	L_3	L_2	L_1	M	U_1	U_2	U_3
+2	L_3	L_2	M	U_2	U_3	U_3	U_3
+1	L_3	L_3	L_1	U_1	U_3	U_3	U_3
0	L_3	L_3	L_2	M	U_2	U_3	U_3
−1	L_3	L_3	L_3	L_1	U_1	U_3	U_3
−2	L_3	L_3	L_3	L_2	M	U_2	U_3



Before considering the output let us interpret the states: M occurs when the streams presently describe the same number; the U-states occur when the upper stream presently describes a number larger than the lower; U_1 when it is possible that the upper stream will end up smaller;[185] U_2 and U_3 when it is known that the upper stream will henceforth always describe a larger number; U_2 when it is possible that the numbers will converge, that is, even though the upper will always be larger the difference may go to zero; U_3 when it is known that the numbers will not converge, that is, the difference is bounded away from zero. For the Ls just replace, obviously, the *upper*s with *lower*s.

For the "max" operation define the output so that it echos the upper digit whenever moving to or from a U-state and echos the lower digit whenever moving to or from an L-state. Note that the only times when the automaton stays in state M is when the upper and lower digits are equal—in that case echo that unique digit. (There is no direct motion between the L- and U-states.) For the "min" operation just reverse, obviously, these output rules.

For the lattice operations one can conflate the outer pairs of states, that is, we can replace every subscript 3 with the subscript 2 to obtain a five-state machine.

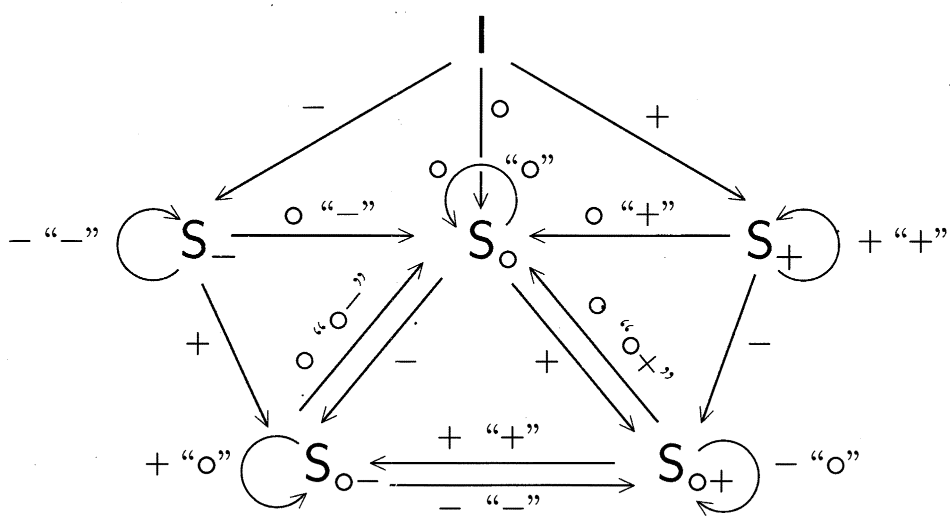
But if we wish for a machine that can tell us when the two streams are describing necessarily unequal numbers we need at least six states. It is possible in states U_2 and L_2 that the numbers are equal (*e.g.,* in state U_2 when all subsequent upper digits are -1 and lower digits $+1$). The six states required for an "apartness" machine can be realized by conflating L_3 and U_3. The seven-state machine will do both tasks: the lattice operations and the apartness information.

[185] If we view the automaton as a Markov process with two absorbing states and if we take the digits as randomly equidistributed then the odds are 13 to 1 that the machine will eventually reach U_3 starting in U_1. Using, instead, the distribution suggested by the "better-stream" automaton (see below), to wit, a 1/2 probability of o and 1/4 for each of + and − the odds are 23 to 1. The expected waiting times are left, of course, as an exercise.

As with the midpoint machine, the lattice machine does not create bad tails: if it were then at least one of the input streams would have to converge to a dyadic rational different from either $\top$ or $\bot$ but by assumption any such input would have to end with all os. [186]

The contrapuntal procedure produces only OK streams and the set of OK streams is closed under action by the zoom, midpoint and lattice machines. But ultimately there will be a need for a finite automaton that converts any stream into an OK Stream. Actually we are handed one, to wit, the result, essentially, of feeding a signed-binary stream into the contrapuntal procedure. But it can be much simplified,

The **OK-stream automaton**:

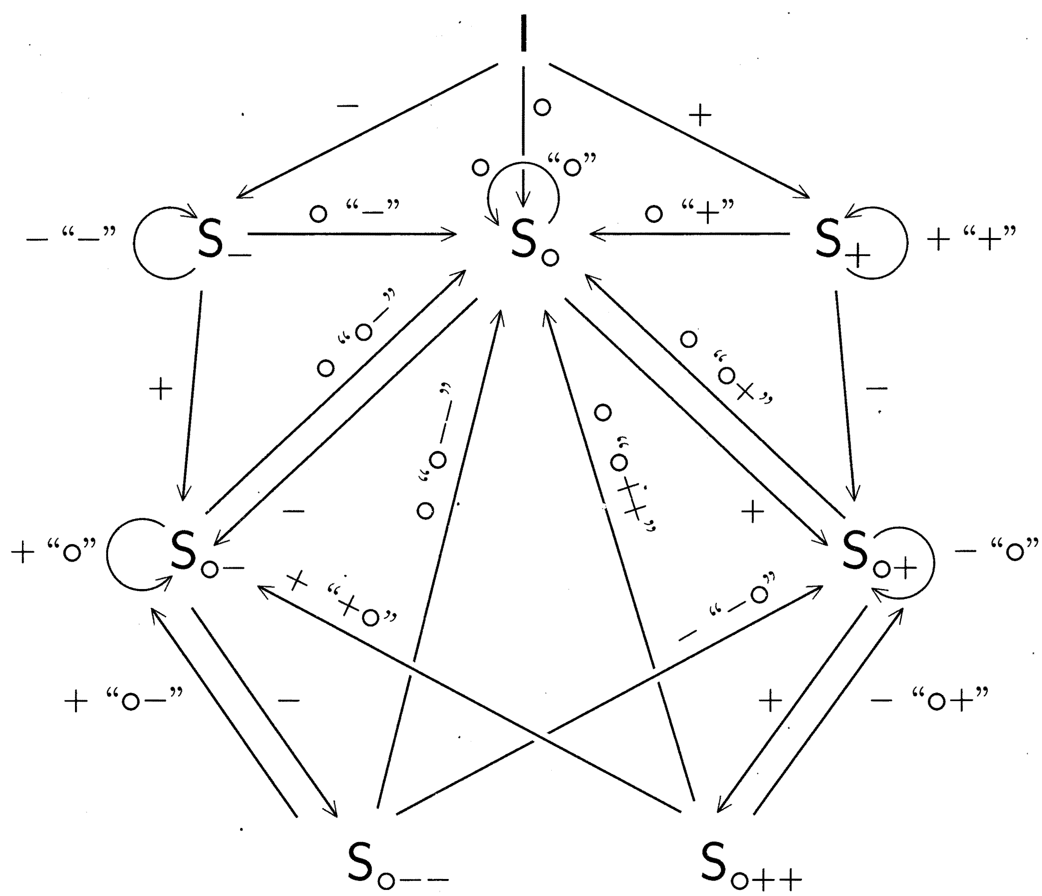


All down-slopping arrows are stammers, all up-slopping arrows are stutters. The two states S_-, S_+ occur only if all inputs have been the same non-zero digit, hence can not be producing a bad tail. It is easy to check that starting from any other state, an input triple of just +s or just −s will always produce an output that includes a o. That the output stream is numerically equivalent to the input stream can be checked by noting that at each step the present input word is numerically equivalent to the result of catenating the present output-word with the word that appears as the subscript in the name of the state (where, of course, the subscript of I is the empty word).

We defined a good stream to be one that described a path that hit an infinity of monodes (on the Houdini diagram). The stream-docking automata of the last section does not always produce good streams. (The stream $+-+-+-\cdots$ is an example of such) To convert any stream into a good stream we could follow a OK stream machine with the machine on page 68.

[186] If we are confident that the input streams are without bad tails we can get by with just 5 states: the states with subscript 2 can be merged with the those with subscript 3. Such can yield quicker apartness results. The words $-+^k$ and o^k- are numerically equal but when we know that there are no bad tails apartness from 0 can be inferred $k+1$ digits earlier with the first than with the second.

Better, use this **good-stream automaton** that does both:



As for the earlier machine, all down-slopping arrows are stammers, all up-sloping arrows are stutters. The only outputs that don't include a $\circ$ are from just two states, S_-, S_+, and those two states occur only on the edges. All outputs, therefore, include a monode. The same argument as previously yields that the input and output streams are numerically equal.

The array below describes a stutter-free version of the good stream automaton. The 28 states are named by words of signed binary digits of length of 0,1,2 and 3. (The initial state is named by the empty word.) They appear in the leftmost column. The input digits appear on the top row. For the 11 transient states (that is, those named by words of length less than 3) each entry describes the next state. For the 17 recurring states named by words of length 3, each entry describes, first, the output digit (between the quotation marks), second the next state. Each entry has the property that the catenation of the output symbol with the "goto-word" is numerically equal to the catenation of the name of the present state with the input.

The **stutter-free good-stream automaton**:

	−	o	+
	−	o	+
+	o+	+o	++
o	o−	oo	o+
−	−−	−o	o−
++	+o+	++o	+++
+o	+o−	+oo	+o+
o+	oo+	o+o	o++
oo	oo−	ooo	oo+
o−	o−−	o+o	o++
−o	−o−	−oo	−o+
−−	−−−	−−o	−o−

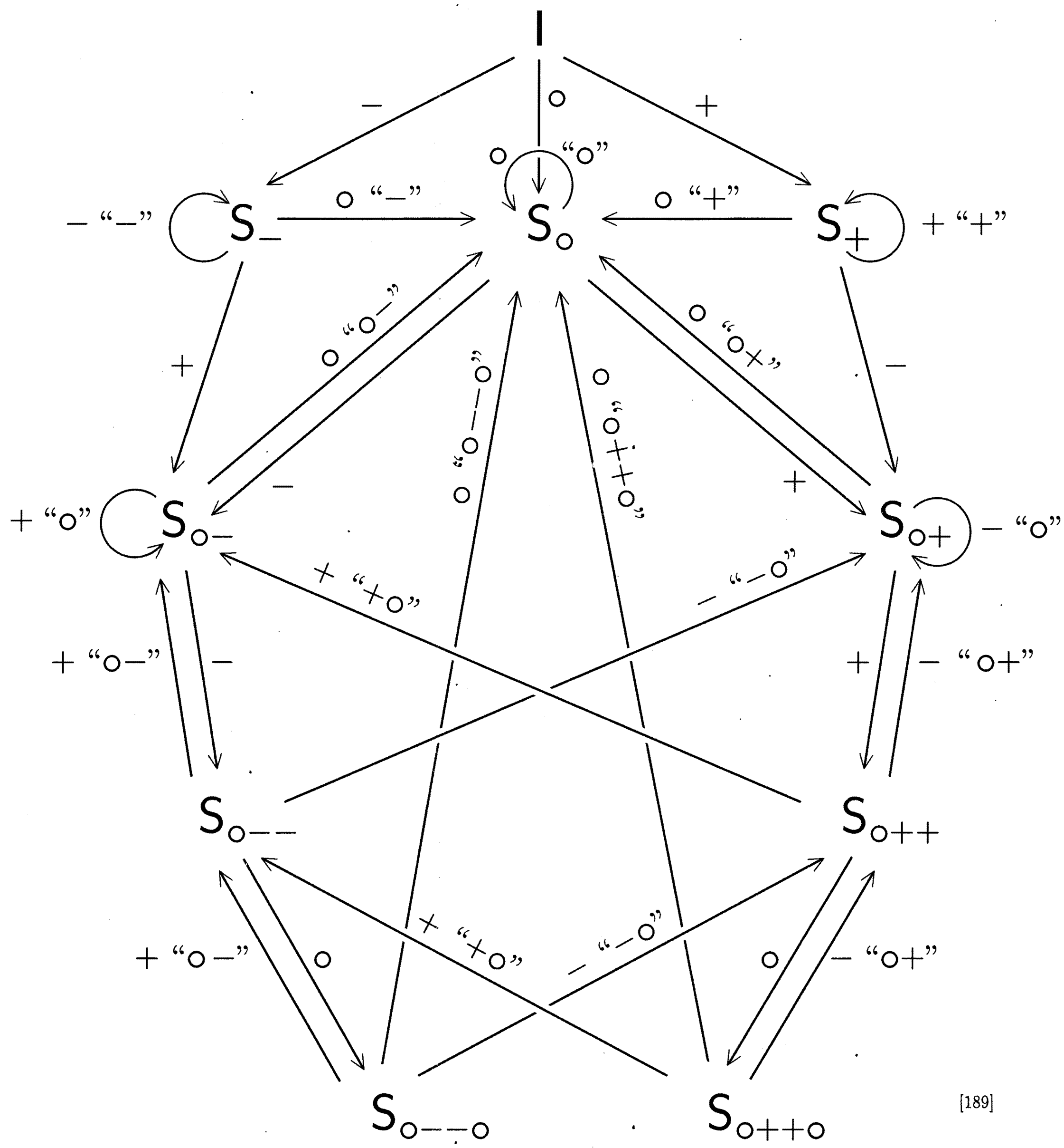
	−		o		+	
+++	"+"	+o+	"+"	++o	"+"	+++
++o	"+"	+o−	"+"	+oo	"+"	+o+
+o+	"+"	oo+	"+"	o+o	"+"	o++
+oo	"+"	oo−	"+"	ooo	"+"	oo+
+o−	"+"	o−−	"+"	o−o	"+"	o−+
o++	"o"	+o+	"o"	++o	"+"	oo−
o+o	"o"	+o−	"o"	+oo	"o"	+o+
oo+	"o"	oo+	"o"	o+o	"o"	o++
ooo	"o"	oo−	"o"	ooo	"o"	oo+
oo−	"o"	o−−	"o"	o−o	"o"	oo−
o−o	"o"	−o−	"o"	−oo	"o"	−o+
o−−	"−"	oo+	"o"	−−o	"o"	−o−
−o+	"−"	oo+	"−"	o+o	"−"	o++
−oo	"−"	oo−	"−"	ooo	"−"	oo+
−o−	"−"	o−−	"−"	o−o	"−"	oo−
−−o	"−"	−o−	"−"	−oo	"−"	−o+
−−−	"−"	−−−	"−"	−−o	"−"	−o−

But there's a better automaton. By adding two more states we obtain the machine shown on the next page, the output of which is asymptotically at least half monode. It suffices to note that each "output word" (a word that's enclosed in quotation marks) is at least half monode.[187] Indeed, the output is not just asymptotically monodal, at least half of any even-length initial segment is monodal.[188] (If any initial word is a catenation of output-words we're done. If not, then it's a catenation of goto-words followed by an initial segment of one of the output-words. But since each even-length initial segment of an output-word is half monodal, we need worry only about a catenation of output-words followed by an initial output-word segment of odd length. In all such initial segments the number of non-monodes is at most one more than the number of monodes. The hypothesis—the assumption that we're looking at an initial segment of even length—forces at least one of the previous output-words to be of length one. And any such word consists of a single monode.)

[187] To wit, the two single-digit edge-outputs. plus o, o−, o+, +o, −o, o−−o, o++o.

[188] Hence for any initial segment of length n the number of monodes is at least the integer part of $n/2$.

The **better-stream automaton**:



[189]

Can we do better with an even bigger machine? Not if the goal is always asymptotically <u>better</u> than half monodal. A fairly simple argument shows that there can not be any pair of adjacent os in any stream of signed binary digits that converges to the number 1/3.[190]

---

[189] Honest, I didn't notice its resemblance to the bejeweled—but no longer worn—symbol of power until several days after its creation.

[190] Start with the fact that any stream starting with oo converges to an element of I between $-1/4$ and $1/4$: if oo were to appear in a stream converging to $1/3$ let n be the number of digits prior to that oo, m the integer (necessarily positive) named by those digits and x the element in the interval from $-1/4$ to $1/4$ named by the stream that remains after those first n digits are removed; if $x > 0$ we could infer from $2^n/3 = m + x$ that x is the "fractional part" of $2^n/3$, that is, it would have to be $1/3$ or $2/3$; but if $x < 0$ then we could infer from $2^n/3 + 1 = m + (1 + x)$ that $1 + x$ is the fractional part of $(2^n + 3)/3$, that is, x would have to be $-1/3$ or $-2/3$. (Note that arbitrarily long initial segments of $+ \circ - \circ - \circ - \cdots$ are less than half monodal; hence the word "asymptotically.")

The **stutter-free better-stream automaton**:

	−	o	+
	−	o	+
+	o+	+̇o	+̇+
o	o−	ȯo	o+
−	−̇−	−̇o	o−
+̇+	+̇o+	+̇+̇o	+̇+̇+
+̇o	+̇o−	+̇ȯo	+̇o+
ȯ+	ȯo+	ȯ+̇o	o++
ȯo	ȯo−	ȯȯo	ȯo+
ȯ−	ȯ−−	ȯ−̇o	ȯo−
−̇o	−̇o−	−̇ȯo	−̇o+
−̇−	−̇−̇−	−̇−̇o	−̇o−
+̇+̇+	+̇+̇o+	+̇+̇+̇o	+̇+̇+̇+
+̇+̇o	+̇+̇o−	+̇+̇ȯo	+̇+̇o+
+̇o+	+̇ȯo+	+̇ȯ+̇o	+̇o++
+̇ȯo	+̇ȯo−	+̇ȯȯo	+̇ȯo+
+̇o−	+̇ȯ−−	+̇ȯ−̇o	+̇ȯȯ−̇
o++	ȯ+̇o+	o++o	+̇ȯo−
ȯ+̇o	ȯ+̇o−	ȯ+̇ȯo	ȯ+̇o+
ȯȯ+	ȯȯo+	ȯȯ+̇o	ȯȯ++
ȯȯo	ȯȯo−	ȯȯȯo	ȯȯo+
ȯo−	ȯȯ−−	ȯȯ−̇o	ȯȯo−
ȯ−̇o	ȯ−̇o−	ȯ−̇ȯo	ȯ−̇o+
ȯ−−	−̇ȯo+	o−−o	ȯ−̇o−
−̇o+	−̇ȯo+	−̇ȯ+̇o	−̇o++
−̇ȯo	−̇ȯo−	−̇ȯȯo	−̇ȯo+
−̇o−	−̇o−−	−̇ȯ−̇o	−̇ȯo−
−̇−̇o	−̇−̇o−	−̇−̇ȯo	−̇−̇o+
−̇−̇−	−̇−̇−̇−	−̇−̇−̇o	−̇−̇o−

	−		o		+	
+̇+̇+̇+	“+”	+̇+̇o+	“+”	+̇+̇+̇o	“+”	+̇+̇+̇+
+̇+̇+̇o	“+”	+̇+̇o−	“+”	+̇+̇ȯo	“+”	+̇+̇o+
+̇+̇o+	“+”	+̇ȯo+	“+”	+̇ȯ+̇o	“+”	+̇o++
+̇+̇ȯo	“+”	+̇ȯo−	“+”	+̇ȯȯo	“+”	+̇ȯo+
+̇+̇o−	“+”	+̇o−−	“+”	+̇ȯ−̇o	“+”	+̇ȯo−
+̇o++	“+”	ȯ+̇o+	“+”	o++o	“+”	+̇ȯo−
+̇ȯ+̇o	“+”	ȯ+̇o−	“+”	ȯ+̇ȯo	“+”	ȯ+̇o+
+̇ȯȯ+	“+”	ȯȯo+	“+”	ȯȯ+̇o	“+”	ȯo++
+̇ȯȯo	“+”	ȯȯo−	“+”	ȯȯȯo	“+”	ȯȯo+
+̇ȯȯ−	“+”	ȯo−−	“+”	ȯȯ−̇o	“+”	ȯȯo−
+̇ȯ−̇o	“+”	ȯ−̇o−	“+”	ȯ−̇ȯo	“+”	ȯ−̇o+
+̇o−−	“+”	−̇ȯo+	“+”	o−−o	“+”	ȯ−̇o−
o++o	“+”	ȯ−̇o−	“o”	+̇+̇ȯo	“+̇”	ȯo−−
ȯ+̇o+	“o”	+̇ȯo+	“o”	+̇ȯ+̇o	“o”	+̇o++
ȯ+̇ȯo	“o”	+̇ȯo−	“o”	+̇ȯȯo	“o”	+̇ȯo+
ȯ+̇o−	“o”	+̇o−−	“o”	+̇ȯ−̇o	“o”	+̇ȯo−
ȯo++	“o”	ȯ+̇o+	“o”	o++o	“o”	+̇ȯo−
ȯȯ+̇o	“o”	ȯ+̇o−	“o”	ȯ+̇ȯo	“o”	ȯ+̇o+
ȯȯȯ+	“o”	ȯȯo+	“o”	ȯȯ+̇o	“o”	ȯo++
ȯȯȯo	“o”	ȯȯo−	“o”	ȯȯȯo	“o”	ȯȯo+
ȯȯȯ−	“o”	ȯo−−	“o”	ȯȯ−̇o	“o”	ȯȯo−
ȯȯ−̇o	“o”	ȯ−̇o−	“o”	ȯ−̇ȯo	“o”	ȯ−̇o+
ȯȯ−−	“o”	−̇ȯo+	“o”	o−−o	“o”	ȯ−̇o−
ȯ−̇o+	“o”	−̇ȯo+	“o”	−̇ȯ+̇o	“o”	−̇o++
ȯ−̇ȯo	“o”	−̇ȯo−	“o”	−̇ȯȯo	“o”	−̇ȯo+
ȯ−̇o−	“o”	−̇o−−	“o”	−̇ȯ−ȯ	“o”	−̇ȯo−
o−−o	“−”	ȯo++	“o”	−̇−̇ȯo	“−”	ȯ+̇o+
−̇o++	“−”	ȯ+̇o+	“−”	o++o	“−”	+̇ȯo−
−̇ȯ+̇o	“−”	ȯ+̇o−	“−”	ȯ+̇ȯo	“−”	ȯ+̇o+
−̇ȯȯ+	“−”	ȯȯo+	“−”	ȯȯ+̇o	“−”	ȯo++
−̇ȯȯo	“−”	ȯȯo−	“−”	ȯȯȯo	“−”	ȯȯo+
−̇ȯȯ−	“−”	ȯo−−	“−”	ȯȯ−̇o	“−”	ȯȯo−
−̇ȯ−̇o	“−”	ȯ−̇o−	“−”	ȯ−̇ȯo	“−”	ȯ−̇o+
−̇o−−	“−”	−̇ȯo+	“−”	o−−o	“−”	ȯ−̇o−
−̇−̇o+	“−”	−̇ȯo+	“−”	−̇ȯ+̇o	“−”	−̇o++
−̇−̇ȯo	“−”	−̇ȯo−	“−”	−̇ȯȯo	“−”	−̇ȯo+
−̇−̇o−	“−”	−̇o−−	“−”	−̇ȯ−̇o	“−”	−̇ȯo+
−̇−̇−̇o	“−”	−̇−̇o−	“−”	−̇−̇ȯo	“−”	−̇−̇o+
−̇−̇−̇−	“−”	−̇−̇−̇−	“−”	−̇−̇−̇o	“−”	−̇−̇o−

The dots have no effect on the behavior of the machine; their purpose is only to ease a (forthcoming) verification that the stutter-free machine's output streams are the same as the those of the diagrammed machine. The 67 states are named by words of signed binary digits of length 0 through 4. (The initial state is named by the empty word.) They appear in the leftmost columns. The input digits appear on the top row. For the 28 transient states (that is, those named by words of length less than 4) each entry describes the next state. For the 39 recurring states (named by words of length 4) each entry describes, first, the output digit (between the quotation marks), second the next state. Each entry has the property that the catenation of the name of the present state with the input symbol is numerically equal to the catenation of the output symbol with the name of the goto-state.

It's another thing to show that it produces the same output stream as the diagrammed machine. That's what the dots are for. They identify those digits guaranteed to be output: if the name of a present state begins with $\dot{x}$ then the output—regardless of input—will be x; the subsequent digits of the name marked with a dot will appear as an initial segment of the next state—again, regardless of input). The rest of digits in the present-state name describe the (subscript) of the state in the diagrammed machine that the stutter-free automaton is emulating. And the rest of the digits in the goto-word describe the next state in the stutter-free machine.[191]

[191] In case you wish to check some of these claims (including the implicit claim that every goto-state is, indeed, one of the

There is yet one more improvement to make. As noted there may be no improvement in the frequency of monodes but at least we can eliminate the 2-digit subwords +− and −+.

One way of course would be to follow the better stream-automaton with a fairly simple automaton designed just for this purpose. But we can, in fact, get by with a fairly minimal adjustment to obtain the stutter-free **even-better-stream automaton**:

	−	o	+
	−	o	+
+	o+	+o	++
o	o−	oo	o+
−	−−	−o	o−
++	+o+	++o	+++
+o	+o−	+oo	+o+
o+	oo+	o+o	o++
oo	oo−	ooo	oo+
o−	o−−	o−o	oo−
−o	−o−	−oo	−o+
−−	−−−	−−o	−o−
+++	++o+	+++o	++++
++o	++o−	++oo	++o+
+o+	+oo+	+o+o	+o++
+oo	+oo−	+ooo	+oo+
+o−	+o−−	+o−o	+oo−
o++	o+o+	o++o	+oo−
o+o	o+o−	o+oo	o+o+
oo+	ooo+	oo+o	oo++
ooo	ooo−	oooo	ooo+
oo−	oo−−	oo−o	ooo−
o−o	o−o−	o−oo	o−o+
o−−	−oo+	o−−o	o−o−
−o+	−oo+	−o+o	−o++
−oo	−oo−	−ooo	−oo+
−o−	−o−−	−o−o	−oo−
−−o	−−o−	−−oo	−−o+
−−−	−−−−	−−−o	−−o−

	−		o		+	
++++	"+"	++o+	"+"	+++o	"+"	++++
+++o	"+"	++o−	"+"	++oo	"+"	++o+
++o+	"+"	+oo+	"+"	+o+o	"+"	+o++
++oo	"+"	+oo−	"+"	+ooo	"+"	+oo+
++o−	"+"	+o−−	"+"	+o−o	"+"	+oo−
+o++	"+"	o+o+	"+"	o++o	"+"	+oo−
+o+o	"+"	o+o−	"+"	o+oo	"+"	o+o+
+oo+	"+"	ooo+	"+"	oo+o	"+"	oo++
+ooo	"+"	ooo−	"+"	oooo	"+"	ooo+
+oo−	"+"	oo−−	"+"	oo−o	"+"	ooo−
+o−o	"+"	o−o−	"+"	o−oo	"+"	o−o+
+o−−	**"o"**	**+oo+**	**"o"**	**+o+o**	"+"	o−o−
o++o	"+"	o−o−	"o"	++oo	"+"	oo−−
o+o+	"o"	+oo+	"o"	+o+o	"o"	+o++
o+oo	"o"	+oo−	"o"	+ooo	"o"	+oo+
o+o−	"o"	+o−−	"o"	+o−o	"o"	+oo−
oo++	"o"	o+o+	"o"	o++o	"o"	+oo−
oo+o	"o"	o+o−	"o"	o+oo	"o"	o+o+
ooo+	"o"	ooo+	"o"	oo+o	"o"	oo++
oooo	"o"	ooo−	"o"	oooo	"o"	ooo+
ooo−	"o"	oo−−	"o"	oo−o	"o"	ooo−
oo−o	"o"	o−o−	"o"	o−oo	"o"	o−o+
oo−−	"o"	−oo+	"o"	o−−o	"o"	o−o−
o−o+	"o"	−oo+	"o"	−o+o	"o"	−o++
o−oo	"o"	−oo−	"o"	−ooo	"o"	−oo+
o−o−	"o"	−o−−	"o"	−o−o	"o"	−oo−
o−−o	"−"	oo++	"o"	−−oo	"−"	o+o+
−o++	"−"	o+o+	**"o"**	**−o−o**	**"o"**	**−oo−**
−o+o	"−"	o+o−	"−"	o+oo	"−"	o+o+
−oo+	"−"	ooo+	"−"	oo+o	"−"	oo++
−ooo	"−"	ooo−	"−"	oooo	"−"	ooo+
−oo−	"−"	oo−−	"−"	oo−o	"−"	ooo−
−o−o	"−"	o−o−	"−"	o−oo	"−"	o−o+
−o−−	"−"	−oo+	"−"	o−−o	"−"	o−o−
−−o+	"−"	−oo+	"−"	−o+o	"−"	−o++
−−oo	"−"	−oo−	"−"	−ooo	"−"	−oo+
−−o−	"−"	−o−−	"−"	−o−o	"−"	−oo+
−−−o	"−"	−−o−	"−"	−−oo	"−"	−−o+
−−−−	"−"	−−−−	"−"	−−−o	"−"	−−o−

The bold face has no effect on the behavior. Its purpose is only to mark the changes from the previous array. In that array there are just two places where a +− is produced, the first is immediately clear in the first of the three responses in the row:

+̇o−− ‖ "+" −̇ȯo+ | "+" o−−o | "+" ȯ−̇o−

The second, not so immediately clear, starts also in state +̇o−−. Starting there the input string o− produces the output + − . [192]

We'll replace the row with the numerically equivalent:

+o−− ‖ "o" +oo+ | "o" +o+o | "+" o−o−

states) you might want to use `http://www.math.upenn.edu/~pjf/analysisCHECK.pdf`

[192] We have now found all possibilities for the production of the output string +−, that is we have found all appearances of "+" followed by a goto-state that starts (case 1) with −̇ or (case 2) with an undotted digit but which allows a "−" among its output possibilities. Each case appears just once following "+" and the two appearances are in the same row.

It strikes me as remarkable that any stream is numerically equal to one consisting of repeated monopodes alternating with the words $+, ++, -$ or $--$. Put another way, all we need are streams of extended Kleene form

$$\{+^*, -^*\}\left\{ \{\circ, \circ+, +\circ, \circ-, -\circ\}^{\mathbb{N}} \cap \{\circ, \{++, +, --, -\}\circ\}^{\mathbb{N}} \right\}$$

These automata can be easily modified for use in the usual real-number context. The first modification inserts a binary point [193] in the proper place in the output stream.[194] The second modification is remarkably simple: add a ○ as the first digit. We don't need the word "monode" (nor $\{+^*, -^*\}$ in the Kleene expression); each even initial segment of the output will be at most one digit away from being at least half ○s.

We can simplify the automaton by adding not just one ○ to the beginning but four ○s. We use 31 states to obtain:

The real-number even-better-stream automaton

	−		○		+	
++○○	"+"	+○○−	"+"	+○○○	"+"	+○○+
+○++	"+"	○+○+	"+"	○++○	"+"	+○○−
+○+○	"+"	○+○−	"+"	○+○○	"+"	○+○+
+○○+	"+"	○○○+	"+"	○○+○	"+"	○○++
+○○○	"+"	○○○−	"+"	○○○○	"+"	○○○+
+○○−	"+"	○○−−	"+"	○○−○	"+"	○○○−
+○−○	"+"	○−○−	"+"	○−○○	"+"	○−○+
+○−−	"○"	+○○+	"○"	+○+○	"+"	○−○−
○++○	"+"	○−○−	"○"	++○○	"+"	○○−−
○+○+	"○"	+○○+	"○"	+○+○	"○"	+○++
○+○○	"○"	+○○−	"○"	+○○○	"○"	+○○+
○+○−	"○"	+○−−	"○"	+○−○	"○"	+○○−
○○++	"○"	○+○+	"○"	○++○	"○"	+○○−
○○+○	"○"	○+○−	"○"	○+○○	"○"	○+○+
○○○+	"○"	○○○+	"○"	○○+○	"○"	○○++
INITIAL STATE → ○○○○	"○"	○○○−	"○"	○○○○	"○"	○○○+
○○○−	"○"	○○−−	"○"	○○−○	"○"	○○○−
○○−○	"○"	○−○−	"○"	○−○○	"○"	○−○+
○○−−	"○"	−○○+	"○"	○−−○	"○"	○−○−
○−○+	"○"	−○○+	"○"	−○+○	"○"	−○++
○−○○	"○"	−○○−	"○"	−○○○	"○"	−○○+
○−○−	"○"	−○−−	"○"	−○−○	"○"	−○○−
○−−○	"−"	○○++	"○"	−−○○	"−"	○+○+
−○++	"−"	○+○+	"○"	−○−○	"○"	−○○−
−○+○	"−"	○+○−	"−"	○+○○	"−"	○+○+
−○○+	"−"	○○○+	"−"	○○+○	"−"	○○++
−○○○	"−"	○○○−	"−"	○○○○	"−"	○○○+
−○○−	"−"	○○−−	"−"	○○−○	"−"	○○○−
−○−○	"−"	○−○−	"−"	○−○○	"−"	○−○+
−○−−	"−"	−○○+	"−"	○−−○	"−"	○−○−
−−○○	"−"	−○○−	"−"	−○○○	"−"	−○○+

Again, each entry has the property that the catenation of the name of the present state with the input symbol is numerically equal to the catenation of the output symbol with the name of the goto-state.

The subwords "+−" and "−+" do not appear and no subword of length four has fewer than two ○s (hence any initial segment is more than half ○).

[193] in lieu of a decimal point

[194] There are two ways of defining *proper place*, one for midpointing, one for addition.

It should be noted that there is a down side in using these automata: they can throw away information. When we know that a stream can not have a bad tail a long word of the form $-++++\cdots$ would be replaced by one that starts with $\text{ooooo}\cdots$ and that delays the information that the stream will not end positively. (If the stream is the output of the contrapuntal procedure then we're throwing away even more: the first word is produced only when the stream is destined to end *negatively*.) In [186] (p100) we pointed out that $-+^k$ can be better than o^k-.

We find ourselves in a situation quite familiar to those born early enough to remember "table of values" for functions such as logs and sines. They were not values, of course, but approximations of values. If the last digit is a 5 and you wish to drop that last digit by further "rounding" you do not know whether to round up or down. Thus many tables used different 5s to indicate whether the 5, itself, was the result of rounding up or down.

For us, perhaps, ◡ and ◠ .

44. *Addendum:* The Rimsky Scale and Dedekind Incisions [2015–1–13]

In this and the next section we'll be working with partial functions. We'll use $=$ signs between partial functions only in the case of "Kleene equality," that is, only when they're equal and have the same domain. (Hence the usual formula for the linearity of derivatives is not a Kleene equality.) We'll use the "venturi tube" $\succeq$ for a "semi-Kleene equality": if the left side is defined then so is the right and the values are equal. (Thus in calculus $f' + g' \succeq (f+g)'$.) [195] When working with sets of partial functions it's often useful to describe their partial ordering by $\succeq$ as "graph-inclusion." We're interested in the connected components of the set of bounded continuous partial functions on dense domains (recall that a poset is "connected" if for every pair of elements there's a sequence of comparable elements connecting them). In fact, each component has a maximal element and we'll call such a **pre-continuous function**, to wit, a continuous bounded partial function with a <u>maximal domain,</u> that is, a domain that cannot be enlarged in a way that maintains continuity. Necessarily such a domain must be dense and without any "removable singularities."

We'll easily show that much of what we'll be doing with scales can be generalized to real Banach algebras. A preview: let $\text{PCF}(X)$ be the Banach algebra whose elements are pre-continuous real-valued functions on X.

What? Banach algebra? We'll prove an important lemma: *If X is a compact metric space and f is a bounded continuous partial function with a maximal domain on X then the domain of f contains a dense G_δ subset of X.*

Yes, it is a Banach algebra. The norm of $f \in \text{PCF}(X)$ is the uniform norm (note that f was required to be bounded). $f+g=h$ and $fg=h$ mean that the equations hold on the intersection of the domains. To construct the limit of a Cauchy sequence $\{f_n\}$ start by working on the restriction of the fs to the intersection of their domains, construct the Cauchy limit thereon and finish by removing all removable singularities.

If I'm the first to describe such a Banach algebra I'm not the first to live in one. I wrote in Section 32 (p75) with regard to scene analysis: "It would be a mistake to restrict to continuous functions—sharp contour boundaries surely want to exist." But long before scene analysis we've been idealizing physical objects as closed subsets of space—or maybe open subsets. We

[195] Note that "Kleene semi-equality" (as opposed to "semi-Kleene equality") would refer—instead—to an ordering relation between a pair of partial functions with the same domain).

tried to ignore the ambiguity.[196]

Define $g^{\downarrow}$ and $g^{\uparrow}$ by

$$g^{\downarrow}(x) \;=\; g(x) \wedge \liminf_{y\to x} g(y) \qquad \text{and} \qquad g^{\uparrow}(x) \;=\; g(x) \vee \limsup_{y\to x} g(y)$$ [197]

Note that $g^{\downarrow}$ is lower-semicontinuous and $g^{\uparrow}$ is upper-semicontinuous,[198] indeed $g^{\downarrow}$ is the highest lower-semicontinuous function below g and $g^{\uparrow}$ is the lowest upper-semicontinuous function above g. One consequence is that the operations denoted by $^{\downarrow}$ and $^{\uparrow}$ are clearly idempotent and covariant.

We say that $\langle \ell, u\rangle$ is a **matched pair of semicontinuous functions** if $\ell^{\uparrow} = u$ and $u^{\downarrow} = \ell$.

Note well: $g^{\downarrow}$ and $g^{\uparrow}$ needn't be matched. A revealing example: let $S \subseteq X$ and $g = \chi_S$. Then $g^{\downarrow}$ is the characteristic function of the interior of S and $g^{\uparrow}$ of its closure. When S is open $g^{\uparrow\downarrow} = g^{\downarrow}$ means that S is a regular open set.

For any $g : X \to \mathrm{I}$ we obtain matched pairs $\langle g^{\uparrow\downarrow}, g^{\uparrow\downarrow\uparrow}\rangle$ and $\langle g^{\downarrow\uparrow\downarrow}, g^{\downarrow\uparrow}\rangle$ and—all together—up to seven functions: $g, g^{\downarrow}, g^{\uparrow}, g^{\downarrow\uparrow}, g^{\uparrow\downarrow}, g^{\downarrow\uparrow\downarrow}, g^{\uparrow\downarrow\uparrow}$. [199]

For a proof it helps to move to a general setting. On an arbitrary poset let $^{\downarrow}$ and $^{\uparrow}$ be a pair of idempotent covariant endofunctions such that $x^{\downarrow} \leq x^{\uparrow}$ for all x. Then we need to show that the semigroup generated by $^{\downarrow}, ^{\uparrow}$ has as at most 6 elements $^{\downarrow}, ^{\uparrow}, ^{\downarrow\uparrow}, ^{\uparrow\downarrow}, ^{\downarrow\uparrow\downarrow}, ^{\uparrow\downarrow\uparrow}$, each of which is an idempotent. For idempotency use $^{\downarrow\uparrow} = {}^{\downarrow\downarrow\downarrow\uparrow} \leq {}^{\downarrow\uparrow\downarrow\uparrow} \leq {}^{\downarrow\uparrow\uparrow\uparrow} = {}^{\downarrow\uparrow}$ and use the dual argument for $^{\uparrow\downarrow} = {}^{\uparrow\downarrow\uparrow\downarrow}$. For $^{\downarrow\uparrow\downarrow}$ and $^{\uparrow\downarrow\uparrow}$, it's easier: $^{\downarrow\uparrow\downarrow\downarrow\uparrow\downarrow} = {}^{\downarrow\uparrow\downarrow\uparrow\downarrow} = {}^{\downarrow\uparrow\downarrow}$. Given any word on $^{\downarrow}$ and $^{\uparrow}$ use, first, their idempotency to reduce to a word of alternating letters and, second, use the idempotency of $^{\downarrow\uparrow}$ and $^{\uparrow\downarrow}$ to reduce to a word of at most three letters. [200] Now add $x^{\downarrow} \leq x \leq x^{\uparrow}$ and—viewing the poset as a category—note that the set of all elements of the form $x^{\downarrow}$ is coreflective with $x^{\downarrow}$ the coreflection of x and, dually, the set of all elements of the form $x^{\uparrow}$ is reflective with $x^{\uparrow}$ as the reflection.

Note well that the set of elements of the form $x^{\downarrow\uparrow}$ is equal to the set of elements of the form $y^{\uparrow\downarrow\uparrow}$: given x take $y = x^{\downarrow}$; given y take $x = y^{\uparrow}$. We'll call this the set of **matched uppers**. It is not, in general, either reflective or coreflective. None the less:

34.6 LEMMA: *Starting with L, a complete lattice, and a pair of idempotent operations such that $x^{\downarrow} \leq x \leq x^{\uparrow}$ the set of matched uppers, $L^{\downarrow\uparrow}$ is also a complete lattice.*

We're done when we show that $L^{\downarrow\uparrow}$ is a coreflective subset of $L^{\uparrow}$. The coreflection sends $x \in L^{\uparrow}$ to $x^{\downarrow\uparrow}$. We need only show that $x^{\downarrow\uparrow} \leq x$ for all $x \in L^{\uparrow}$. Our assumption is $x^{\downarrow} \leq x$ hence $x^{\downarrow\uparrow} \leq x^{\uparrow}$ and, finally, $x^{\downarrow\uparrow} \leq x^{\uparrow} = x$ for all $x \in L^{\uparrow}$.

Of course the dual argument works for $L^{\uparrow\downarrow}$ the set of matched lowers. But we don't need it. We are handed an isomorphism and and its inverse between $L^{\downarrow\uparrow}$ and $L^{\uparrow\downarrow}$ to wit, the maps denoted by $^{\downarrow}$ and $^{\uparrow}$.

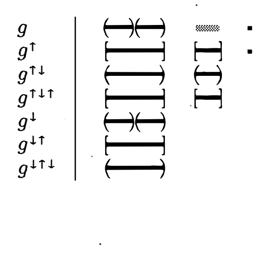
Given any bounded continuous $g : S \to \mathbb{R}$ where S is a dense subset of a space X we obtain a matched pair on X $\langle g^{\downarrow}, g^{\uparrow}\rangle$.

[196] When two objects come in contact are they still disjoint?

[197] Note "$\leq$" and "$\geq$" in the standard definitions of lower- and upper-semicontinuity: "$g(x) \leq \liminf_{y\to x} g(y)$" and "$g(x) \geq \limsup_{y\to x} g(y)$" (the value at x is ignored) hence the longer definition of $g^{\downarrow}$ and $g^{\downarrow}$.

[198] That is, $(g^{\downarrow})^{\downarrow} = g^{\downarrow}$ and $(g^{\uparrow})^{\uparrow} = g^{\uparrow}$.

[199] For an example where all the values are different take g to be the characteristic function of a subset of I consisting, first, of an open subinterval with one point removed, second, a subinterval with irrational elements removed, third, a single isolated point.



For the record, $g^{\downarrow} \leq g^{\downarrow\uparrow\downarrow} \leq h \leq g^{\uparrow\downarrow\uparrow} \leq g^{\uparrow}$ for any g where h can be either $g^{\downarrow\uparrow}$ or $g^{\uparrow\downarrow}$.

[200] Indeed, this sentence shows any idempotent semigroup with two generators has at most 6 elements.

For $x \in S$, of course, $g^{\downarrow}(x) = g(x) = g^{\uparrow}(x)$. For any $x \in X$ there's no problem in seeing that $g^{\downarrow\uparrow}(x) \leqq g^{\uparrow}(x)$. In order to establish that $g^{\downarrow\uparrow}(x) \geqq g^{\uparrow}(x)$ we need to show that for any $\varepsilon > 0$ and any neighborhood of x there exists $y \in S$ such that $g^{\downarrow}(y) \geqq g^{\uparrow}(x) - \varepsilon$. But we know that there exists such a y in S in that neighborhood such that $g(y) \geqq g^{\uparrow}(x) - \varepsilon$, But, as just observed, $g^{\downarrow}(y) = g(y)$. The dual proof yields $g^{\uparrow\downarrow}(x) = g^{\downarrow}(x)$.

Note that if $\langle \ell, u \rangle$ is a matched pair then it's equal to $\langle g^{\downarrow}, g^{\uparrow} \rangle$ for any $\ell \leqq g \leqq u$.

Theorem: *For any matched pair $\langle \ell, u \rangle$ the equalizer of ℓ and u is a dense subset, indeed, the intersection of the graphs of ℓ and u is the graph of a pre-continuous function and all pre-continuous functions so arise. Moreover the domain of any pre-continuous function is* ***substantial,***[201] *that is, it contains a dense G_δ.*

We'll view the equalizer of ℓ and u as the intersection of the sets for all positive ε :

$$S_\varepsilon = \{ \ x \in X \ : \ |u(x) - \ell(x)| < \varepsilon \ \}$$ [202]

We need that S_ε is open and dense.

For any open real interval (a, b) let $\boldsymbol{\langle \ell, u \rangle^{-1}(a, b)} \subseteq X$ denote

$$\{ \ x \in X \ : \ a < \ell(x) \ \text{ and } \ u(x) < b \ \}$$

Because lower semi-continuous maps are precisely the functions whose inverse images carry open updeals to open sets in X and dually for upper semi-continuous maps we have that that $\langle \ell, u \rangle^{-1}(a, b)$ is open. Hence so is

$$S_n = \bigcup_{x \in \mathtt{I}} \langle \ell, u \rangle^{-1}(x, x + \tfrac{1}{n})$$

For the density of S_ε suppose that $U \subseteq X$ is open and disjoint from S_ε. We need to show that U is empty. The characteristic map χ_U is lower semi-continuous, therefore so is $\ell + \varepsilon\chi_U$.[203] Since ℓ is the highest lower semi-continuous map below u it must be that U is empty.

45. *Addendum:* Lebesgue Integration and Measure, Rethought [2015–5–18]

In this section the interval $\mathtt{I}$ will be understood to be the unit interval, $[0, 1]$.

Theories of integration and of measure are, of course, intimately related but they differ in their motivations.

We start with the first, the theory of integration. In Section 40 (p91–94) we described a covariant mean-value function $\mathcal{C}(\mathtt{I}^n) \longrightarrow \mathtt{I}$ that preserves top, bottom and midpointing. We'll denote its values here as $\|\boldsymbol{f}\|_1$ and use it to establish a metric space structure on $\mathcal{C}(\mathtt{I}^n)$.[204] Our goal is to extend this function to a larger scale of **integrable** functions denoted as $\mathcal{L}(\mathtt{I}^n)$. The simplest answer was given by Peter Lax: take $\mathcal{L}(\mathtt{I}^n)$ to be the L^1- metric-space completion of $\mathcal{C}(\mathtt{I}^n)$. [205] That simple answer was likely given by many others. Lax's great

[201] Sometimes "comeagre," sometimes "residual," sometimes "second category."
[202] Yes, the absolute-value bars aren't needed.

[203] The easiest proof uses the view that lower semi-continuity is equivalent to continuity where the target's only open sets are the open downdeals but there's an approach here that avoids using that eqivalence: replace $\varepsilon\chi_U$ with a positive multiple of $dist\langle x, X \setminus U \rangle$. It's easy to see that the sum of a continuous and a semi-continuous is semi-continuous.

[204] As usual, for $f, g \in \mathcal{C}(\mathtt{I}^n)$ we understand $\|f - g\|_1$ to be the mean value of the absolute difference of f, g. Note that $\| \ \|_1$ is quite different from the intrinsic metric (p51), to wit, the one usually denoted $\| \ \|_\infty$.

[205] **Rethinking the Lebesgue integral.** *Amer. Math. Monthly* 116 (2009), no. 10, 863–881. Lax wrote with regard to the other approach (that is, the theory of measure): "In our development Lebesgue measure is a secondary notion. A set S is measurable

contribution was to give us a description of the elements of $\mathcal{L}(\mathbb{I}^n)$ better than "equivalence-classes-of-$\mathbb{L}^1$-Cauchy-sequences": he gave us what he labeled "realizations." And since we're looking at functions that are $\mathbb{I}$-valued (rather than $\mathbb{R}$-valued) it's even easier to describe them. Hold on.

We first need just a little from the other view. Given an open set $U \subseteq \mathbb{I}^n$ define $\mu(U)$ to be the supremum of the mean values of all continuous $h : \mathbb{I}^n \to \mathbb{I}$ with support contained in U:

$$\mu(U) = \sup\{ \ \|h\|_1 \ : \ h \in \mathcal{C}(\mathbb{I}^n), \ \overline{h} \subseteq U \ \}$$

Three important lemmas:

$$\text{If } U_1 \subseteq U_2 \subseteq \cdots \text{ then } \mu(U_1 \cup U_2 \cup \cdots) = \sup_i \mu(U_i)$$ [206]

$$\mu(U_0 \cup U_1)) \leq \mu(U_0) + \mu(U_1)$$ [207]

These two lemmas combine for the third:

$$\mu\left(\bigcup_i U_i\right) \leq \sum_i \mu(U_i)$$

We'll say that a set $S \subseteq \mathbb{I}^m$ is **negligible** if it is contained in open sets of arbitrarily small μ-value and that it is **pervasive** if its complement is negligible. An easy—but necessary to our purposes—simplification: there's no need for the functions to be defined everywhere. We'll allow partial functions whose domains are pervasive.

We'll use the simple fact that a partial function is a special case of a relation, and a relation is—as usual—a set of ordered pairs. We replace Lax's realization with **virtual map**, to wit, a partial function with a pervasive domain whose graph is a countable union of closed subsets, that is, an F_σ set. (Without loss of generality, all F_σs will be understood to be *ascending* unions of closed sets.)

Lax's definition of "realization" was a function f for which there is an $\mathbb{L}^1$-Cauchy sequence of continuous functions that almost everywhere pointwise coverges to f. We'll use virtual maps instead. Our first task is to find such a sequence for any virtual map. And that's easy. Use Tietze [208] to extend each of the continuous partial maps on closed domains to an entire continuous map. The resulting sequence is not only pointwise convergent on the domain of the virtual map: it is *pointwise eventually constant.* Done.

---

if its characteristic function is one of the functions in $\mathbb{L}^1$. Its measure is defined as the integral of the characteristic function. To be sure, such an approach is anathema to probabilists; their object of desire is the "σ-algebra of measurable sets."

[206] It is clear that μ is covariant. For any h and ε where $h \in \mathcal{C}(\mathbb{I}^n)$ has its support in the union of the U_is and ε is positive it suffices to show that $\|h\|_1 \leq \mu(U_n) + \varepsilon$ for large n. Since the compact set $\{ x : h(x) \geq \varepsilon \}$ is contained in the union of the U_is we can find n such that it is contained in U_n. Let $\lfloor h - \varepsilon \rfloor$ be the function such that $\lfloor h - \varepsilon \rfloor(x) = \max\{0, h(x) - \varepsilon\}$. Then the support of $\lfloor h - \varepsilon \rfloor$ is contained in U_n and $\|h\|_1 \leq \|\lfloor h - \varepsilon \rfloor + \varepsilon\|_1 \leq \|\lfloor h - \varepsilon \rfloor\|_1 + \|\varepsilon\|_1 \leq \mu(U_n) + \varepsilon$.

[207] If o $g_0, g_1 \in \mathcal{C}(U_0 \cup U_1)$ are such that $g_0 + g_1$ is constantly 1 and $\overline{g}_i \subseteq U_i$, then for any $h \in \mathcal{C}(\mathbb{I}^n)$ such that $\overline{h} \subseteq U_0 \cup U_1$ we can continuously extend each g_i to all of $\mathbb{I}^n$ by taking 0 as its value off of $U_1 \cup U_2$ thus obtaining $g_0 h + g_1 h = h$ and $\overline{g_i h} \subseteq U_i$. Define $g_i(x) = dist\langle x, \mathbb{I}^n \setminus U_i \rangle \big/ \big(dist\langle x, \mathbb{I}^n \setminus U_0 \rangle + dist\langle x, \mathbb{I}^n \setminus U_1 \rangle \big)$.

[208] The lemma was proved by L.E.J. Brouwer and H. Lebesgue for $\mathbb{R}^n$, by H. Tietze for arbitrary metric spaces, and by P.S. Urysohn for normal spaces. I think it noteworthy that Tietze didn't need the axiom of choice. (Urysohn needed it when he generalized Tietze's theorem to T_4 spaces.) For the record: Given a closed subset A of a metric space X and a continuous $f : A \to [-1, +1]$ define $g_n : X \to [-1, +1]$ inductively by taking g_0 to be constantly 0, and $g_{n+1}x = g_n x + \frac{2^{n-1}}{3^n} \frac{dist\langle x, B_- \rangle - dist\langle x, B_+ \rangle}{dist\langle x, B_- \rangle + dist\langle x, B_+ \rangle}$ where (for $e = \pm$) $B_e = \{x \in A : 2^{n-1}/3^n \leq e(fx - g_n x)\}$ (Use $dist\langle x, \emptyset \rangle = 1$.)

Then $|fx - g_n x| \leq (2/3)^n$ for all $x \in A$. and since $|g_{n+1}x - g_n x| \leq 2^{n-1}/3^n$ for all $x \in X$ the gs uniformly converge to a continuous map.

We'll be using the fact that when its target is compact a function is continuous iff it's graph is a closed set of ordered pairs. [209] Hence a partial $\mathbf{I}$-valued function is F_σ iff it is the union of a sequence of continuous partial functions, each with a closed domain. And it is a virtual map if, moreover, the union of those closed domains is a pervasive set.

F_σ sets are closed, of course, under finite intersection. An easy exercise is that there's a pervasive set on which two virtual maps agree iff their intersection is not only F_σ but still a virtual map, that is, it still has a pervasive domain. We'll abbreviate all that as $f \overset{\text{æ}}{=} g$.

We take the $\overset{\text{æ}}{=}$-classes of virtual maps as the elements of $\mathcal{L}(\mathbf{I}^n)$. [210]

The converse is harder:

45.2 THEOREM: (Peter Lax) *An L^1-Caucy sequence of continuous functions converges to a virtual map.*

Given an L^1-Cauchy sequence $\{f_n\}$ of continuous functions we follow Lax and replace the sequence with a "rapidly converging" subsequence, to wit, one that satisfies the condition $\|f_n - f_{n+1}\|_1 \leqq 1/4^n$. (Working in the metric completion we simply chose the n^{th} entry to be within $1/2^{2n+1}$ of the sequence's limit)[211]

It suffices to find an ascending chain of closed sets $\{A_n\}$ such that the f_ns converge uniformly on A_n for each n and such that $A_1 \cup A_2 \cup \cdots$ is pervasive. First define a sequence of open sets $U_n = \{\, x : 1/2^n < |f_n(x) - f_{n+1}(x)| \,\}$. (Put another way, U_n is the largest open set such that $2^{-n}\chi_{U_n}(x) \leqq |f_n(x) - f_{n+1}(x)|$ for all $x \in \mathbf{I}^n$.) Then $\mu(U_n) \leqq 1/2^n$ (because if $\mu(U_n) > 1/2^n$ there would exist continuous $h \leqq \chi_{U_n}$ such that $1/2^n < \|h\|_1$ but $2^{-n}h < 2^{-n}\chi_{U_n} \leqq |f_n - f_{n+1}|$ would then imply $1/4^n < 2^{-n}\|h\|_1 = \|2^{-n}h\|_1 \leqq \|f_n - f_{n+1}\|_1$).

Define A_n to be the complement of $U_n \cup U_{n+1} \cup U_{n+2} \cup \cdots$. For every $x \in A_n$ it is the case that $|f_j(x) - f_{j+1}(x)| \leqq 1/2^j$ for all $j \geqq n$, hence the f_js converge uniformly on A_n. Finally, the complement of $A_1 \cup A_2 \cup \cdots$ is negligible because for each n it is contained in the complement of A_n and that complement is the union of open sets $U_n, U_{n+1}, U_{n+2}, \ldots$ where the sum $\mu(U_n) + \mu(U_{n+1}) + \mu(U_{n+2}) + \cdots$ is at most $1/2^{n-1}$

(A little modification of the proof works for the L^p-norm for any finite p larger than 1. [212] Given any two L^1-Cauchy sequences of continuous functions that pointwise converge to the same virtual map it's easy to check that the sequence of absolute differences converges to the function that's constantly 0. In particular, their L^1-norms converge to each other and we use that to define the L^1-structure for $\mathcal{L}(\mathbf{I}^n)$.

We extend the meaning of $\overset{\text{æ}}{=}$ to arbitrary functions: $g \overset{\text{æ}}{=} f$ if they agree on a pervasive set. We say that $f : \mathbf{I}^n \dashrightarrow \mathbf{I}$ is a **measurable function** if there is a virtual map g such that $g \overset{\text{æ}}{=} f$.

---

[209] Reaping once again the advantage of $\mathbf{I}$ over $\mathbb{R}$. Note that a "quasi-inverse" function on $\mathbb{R}$ (indeed, any entire function that extends the partial function which sends $x \neq 0$ to x^{-1}) has a closed subset as graph. Much of this material, thought, does generalize to $\mathbb{R}$-valued functions. Just replace F_σ with *countable union of compact subsets.*

[210] When working with partial functions the use of $=$ signs can be misleading. A "Kleene equality" between partial functions says that if either is defined then so is the other and the values are equal. (Hence the usual formula for the linearity of derivatives is not a Kleene equality.) The "venturi tube" $\succcurlyeq$ is used for a "semi-Kleene equality": if the left side is defined then so is the right and the values are equal. (Hence $f' + g' \succcurlyeq (f+g)'$.) The set of virtual maps is partially ordered by $\succcurlyeq$. A poset is "connected" if for every pair of elements there's a sequence of comparable elements connecting them. The elements of $\mathcal{L}(\mathbf{I}^n)$ could be taken to be the connected compontents of the "Kleene-poset" of virtual maps.

[211] The big difference between this proof and Lax's is that his n^2 becomes my 2^n (and his n^4 becomes my 4^n). I don't know why Lax appears to prefer Bernoulli to Zeno (but, alas, looking at n^2 and 2^n I know too well how strephosymbolia would have caused me to so appear).

[212] Note that the rapidly convergent subsequence converges pointwise to the virtual map. The Carleson-Hunt theorem says any subsequence converges pointwise ae. If we take $\mathbb{R}$ as the target, instead of $\mathbf{I}$ the proof continues to work but, as pointed out in footnote [209], the definition of virtual map has to be changed: replace F_σ with "*countable union of compact subsets.*"

Now for the other approach, the theory of Lebesgue measure. Theorem 24.7 (p55) may be viewed as its foundation: any order-complete scale is the injective envelope of its subscale generated by its extreme points. The extreme points in $\mathcal{L}(\mathtt{I}^n)$ are the elements of the form χ_S where $S \subseteq \mathtt{I}^n$. Note that we're not allowing any old subset, just those for which χ_S is measurable, or as usually said, is a **measurable subset**. We write $S_1 \overset{æ}{=} S_2$ when $\chi_{S_1} \overset{æ}{=} \chi_{S_2}$.[213] Easy lemmas tell us that the family of measurable subsets is a σ-algebra, that is, a countably complete Boolean algebra. When we move to the quotient Boolean algebra obtained by identifying $\overset{æ}{=}$-classes we obtain a complete Boolean algebra.[214] Every measurable set is $\overset{æ}{=}$ to an F_σ. Complements of measurable sets are measuable (replace χ_S with $1 - \chi_S$). Thus every measurable set is $\overset{æ}{=}$ to a countable intersection of open sets, that is, a G_δ. In a metric space any open set is a F_σ, quite enough to ensure that its characteristic function has an F_σ graph.

The argument that the characteristic functions in $\mathcal{L}(\mathtt{I}^n)$ form a complete Boolean algebra works as well for all of $\mathcal{L}(\mathtt{I}^n)$. In Section 23 (p51–53) we saw that such means that $\mathcal{L}(\mathtt{I}^n)$ is an injective scale. Alas, it is not the injective envelope of $\mathcal{C}(\mathtt{I}^n)$. Lemma(24.5) (p54) says that such would require $\mathcal{C}(\mathtt{I}^n)_\star$ to be cofinal in $\mathcal{L}(\mathtt{I}^n)_\star$. But if U is a non-pervasive dense open set [215] then there's only one continuous function f such that $\chi_U \leqq f$.

The problem, then, is to find a simply described scale whose injective envelope is $\mathcal{L}(\mathtt{I}^n)$. One solution is the Boolean-algebra scale [216] $\mathbb{I}[B]$ where B is the minimal Boolean algebra that contains a copy of the lattice of open sets in $\mathtt{I}^n$. Start first with the sub-Boolean algebra in the power-set of $\mathtt{I}^n$ generated by the open sets, traditionally called the family of "constructible sets," Such does not satisfy the minimallity condition. So reduce by the ideal of negligiable subsets.

This reduction needs to be looked at carefully. Any construtible set is a finite union of locally closed subsets, the latter being an intersection of an open and a closed.

Let me unravel this. $f \in \mathbb{I}[B]$ if there's a finite partioning of $\mathtt{I}^n$ in which each cell is of the form $U \setminus V$ where U and V are open subsets and f is constantly equal to an element in $\mathbb{I}$ on $U \setminus V$.[217] It's clear that the family of such functions are closed under action by the unary scale operations. For midpointing first take the common refinement of the two partionings (using the equation $(U_1 \setminus V_1) \cap (U_2 \setminus V_2) = (U_1 \cap U_2) \setminus (V_1 \cup V_2)$). The set of such functions is a Boolean-algebra scale. But it is not the minimal such. So reduce by the $\top$-face of all such functions with the property that for each $U \setminus V$ in the partitioning either $f = 1$ or $\mu(U) = \mu(U \cap V)$. Now take the injective envelope of the result.[218]

We need to prove that what we get is $\mathcal{L}(\mathtt{I}^n)$. Since any open set—hence any locally closed set—in a metric space is an F_σ the elements of $\mathbb{I}[B]$ are all virtual maps it suffices to show that for every virtual map f either $f \overset{æ}{=} \top$ or there's a non-empty open $U \subseteq \mathtt{I}^n$ and $m \in \mathbb{N}$ such that $\{\, x : f(x) \leqq (\top|)^m \chi_U(x) \,\}$ is a pervasive set. So let the graph of f be an increasing

[213] Eqivalently when the symmetric difference, $(S_1 \setminus S_2) \cup (S_2 \setminus S_1)$, is negligible.

[214] When I noticed this decades ago, I was surprised. So—amazingly—was everybody I told it to. But it's an easy consequence of countable completeness and the fact that an element of maximal L^1-value in a sub-family closed under countable unions is necessarily that sub-family's maximum element.

[215] Take a countable dense subset and cover its j^{th} element with an open neighborhood of measure $1/4^{j+1}$

[216] Section 38 (p89–91)

[217] To see a nice collection of the graphs of such functions (on $\mathtt{I}^2$) just Google "contour models" and click on "images."

[218] The traditional name for a set of the form $U \setminus V$, where U and V are open is "locally closed set" (alternately, described as the intersecion of an open and a closed, or as a set that's open in its closure). A finite union of locally closed sets is called a "construcitble set." Note that the when we have shown that for two partionings in which each cell is locally closed there's a common refinement with the same property we have shown that each constructible set is a union of pairwise disjoint locally closed sets and, further, has a constructible complement.

union of closed sets $f_1 \cup f_2 \cup f_3 \cup \cdots$. Assuming that f does not name the top element of $\mathcal{L}(\mathtt{I}^n)$ there exists i such that $C = \{\, x : f_i(x) < 1 \,\}$ is of positive measure. For each $m \in \mathbb{N}$ let A_m be the closed set $\{\, x : f(x) \leq 1 - 1/2^m \,\}$. Since $C = \bigcup_m A_m$ there exists m such that $\mu(A_m) > 0$. Define the open set U to be the complement of A_m. Then $f(x) \leq (\top|)^m \chi_U(x)$ on the domain of f and $0 < \|(\top|)^m \chi_U\|_1$.

MORE TO COME [219]

[219] Given countable atomless Boolean algebras we can build a one-to-one isomorphism between them by creating an ascending sequence of finite boolean algebras in each atomless algebra, each finite algebra being equipped with an isomorphism with its corresponding finite algebra, each such isomorphism being an extension of the isomorphism between the previous pair of algebras. Start each sequence with the two-element boolean algebra. Thereafter we follow the "back-and-forth" strategy Cantor used to construct isomorphisms between countable "densely" ordered sets, that is, we alternatingly choose an element in one of the two atomless algebras not yet involved in the correspondence and find an element in the other that will be "similarly situated," that is, will yield an extension of the isomorphism between the two subalgebras that result from the choices. First, note that a pair of finite Boolean algebras are isomorphic iff they have the same number of atoms, moreover, any one-to-one correspondence between their sets of atoms extends uniquely to an isomorphism between the algebras. So when we choose an element not yet in the correspondence it generates a larger finite algebra. Some of the old atoms may be contained in the new element, some may be disjoint from the new element, some may split (into two new atoms) because of the new element. For each such old atom chose a splitting of its corresponding element in the other atomless algebra and obtain two sets of new atoms with a correspondence between them. Now extended that correspondence to an isomorphism between the finite algebras thus generated.

46. *Addendum:* A Few Subscorings [2015–7–27]

It is said that “subscoring” is short for “substitution underscoring,” to wit, a one-column array wherein the underscores indicate the sub-strings to be replaced.[220]

Page 9

$$
\begin{array}{c}
\underline{\underline{a\lhd(a|x)}}\\
(((\dot{a}\underline{\underline{|\bot)|(a}}|x))^{\vee})^{\wedge}\\
(((\underline{\underline{\dot{a}|a}})|(\bot|x))^{\vee})^{\wedge}\\
((\underline{\underline{(\bot|\top)|(\bot|x)}})^{\vee})^{\wedge}\\
(\underline{\underline{(\bot|(\top|x))^{\vee}}})^{\wedge}\\
\underline{\underline{(\top|x)^{\wedge}}}\\
x
\end{array}
$$

Page 10n

$$
\begin{array}{c}
\underline{\underline{x}}|\underline{\underline{\odot}}\\
\underline{\underline{(x|x)|(\dot{x}}}|x)\\
(x|\underline{\underline{\dot{x})}}|\underline{\underline{(x|x)}}\\
\underline{\underline{(\dot{\dot{x}}|\dot{x})}}|x\\
\odot|x
\end{array}
\qquad\qquad
\begin{array}{c}
\underline{\underline{x|y}}\\
\odot\lhd(\underline{\underline{\odot|(x|y)}})\\
\odot\lhd((\odot|x)|(\odot|\underline{\underline{y)}})\\
\odot\lhd((\odot|\underline{\underline{x)|(y}}|\odot))\\
\odot\lhd((\odot|y)|(\underline{\underline{x|\odot}}))\\
\odot\lhd(\underline{\underline{(\odot|y)|(\odot|x)}})\\
\underline{\underline{\odot\lhd(\odot|(y|x))}}\\
y|x
\end{array}
$$

---

[220] The macro {\scor}[1]{\uuline{\rule[-7pt]{0pt}{0pt}#1}} (using package ulem) is useful in their construction.

Page 14–15

$$
\begin{gathered}
(z|x) \multimap (z|y) \\
\hline\hline
((z|x)^{\bullet}|(z|y))^{\vee} \\
((\dot{z}|\dot{x})|(z|y))^{\vee} \\
((\dot{z}|z)|(\dot{x}|y))^{\vee} \\
(\odot|(\dot{x}|y))^{\vee} \\
\hline\hline
\top|(\dot{x}|y)^{\vee} \\
\top|(x \multimap y)
\end{gathered}
$$

$$
\begin{gathered}
(u|v) \multimap (w|x) \\
\hline\hline
((u|v)^{\bullet}|(w|x))^{\vee} \\
((\dot{u}|\dot{v})|(w|x))^{\vee} \\
((\dot{u}|w)|(\dot{v}|x))^{\vee} \\
((u|\dot{w})^{\bullet}|(\dot{v}|x))^{\vee} \\
\hline\hline
(u|\dot{w}) \multimap (\dot{v}|x)
\end{gathered}
$$

Page 23n

$$
\begin{gathered}
\odot\triangleleft(\odot\triangleleft((\dot{a}|\odot)|(a|x))) \\
\odot\triangleleft(\odot\triangleleft((\dot{a}|a)|(\odot|x))) \\
\odot\triangleleft(\odot\triangleleft(\odot|(\odot|x))) \\
\hline\hline
\odot\triangleleft(\odot|x) \\
\hline\hline
x
\end{gathered}
$$

$$
\begin{gathered}
a\triangleleft(a\triangleleft((\odot|a)|(\odot|x)^{\bullet}))^{\bullet} \\
a\triangleleft(a\triangleleft((\odot|a)|(\odot|\dot{x})))^{\bullet} \\
a\triangleleft(a\triangleleft(\odot|(a|\dot{x})))^{\bullet} \\
a\triangleleft(a\triangleleft((a|\dot{a})|(a|\dot{x})))^{\bullet} \\
a\triangleleft(a\triangleleft(a|(\dot{a}|\dot{x})))^{\bullet} \\
\hline\hline
a\triangleleft(\dot{a}|\dot{x})^{\bullet} \\
a\triangleleft(a|x) \\
\hline\hline
x
\end{gathered}
$$

Page 25n

$$x \equiv y$$

$$\begin{array}{cc}
\neg x & \neg y \\
= & = \\
\neg x \wedge \top & \top \wedge \neg y \\
= & = \\
\neg x \wedge (y \vee \neg y) & (\neg x \vee x) \wedge \neg y \\
= & = \\
\neg x \wedge (x \vee \neg y) & (\neg x \vee y) \wedge \neg y \\
= & = \\
(\neg x \wedge x) \vee (\neg x \wedge \neg y) & (\neg x \wedge \neg y) \vee (y \wedge \neg y) \\
= & = \\
\bot \vee (\neg x \wedge \neg y) & (\neg x \wedge \neg y) \vee \bot \\
= & =
\end{array}$$

$$\neg x \wedge \neg y$$

$$x^2 x^* = x = x^{**}$$
$$x x^* = x^* x$$
$$x \equiv y$$

$$\begin{array}{ccc}
x^* & & y^* \\
= & & = \\
x^{*2} x & & y y^{*2} \\
= & & = \\
x^{*2} y & & x y^{*2} \\
= & & = \\
x^{*2} y y y^* & & x^* x x y^{*2} \\
= & & = \\
x^{*2} x y y^* y^* & & x^* x^* x x x y^{*2} \\
= & & = \\
x^{*2} y^3 y^{*2} & \equiv & x^{*2} x^3 y^{*2} \\
= & & =
\end{array}$$

$$u \equiv v$$

$$\begin{array}{cc}
\dot{u} & \dot{v} \\
= & = \\
\odot \triangleleft (\odot | \dot{u}) & \odot \triangleleft (\dot{v} | \odot) \\
= & = \\
\odot \triangleleft ((\dot{v} | v) | \dot{u}) & \odot \triangleleft (\dot{v} | (\dot{u} | u)) \\
= & = \\
\odot \triangleleft ((\dot{v} | u) | \dot{u}) & \odot \triangleleft (\dot{v} | (\dot{u} | v)) \\
= & = \\
\odot \triangleleft ((\dot{v} | \dot{u}) | (u | \dot{u})) & \odot \triangleleft ((\dot{v} | \dot{u}) | (\dot{v} | v)) \\
= & =
\end{array}$$

$$\odot \triangleleft ((\dot{v} | \dot{u}) | \odot)$$

Page 43

$$
\begin{array}{ccc}
e & & \\
= & & \\
e \vee \overline{\bot} & & \\
= & & \\
e \vee \overline{(e \wedge \dot{e})} & & \\
= & & \\
e \vee (\overline{e} \wedge \overline{\dot{e}}) & & \\
= & & \\
(e \vee \overline{e}) \wedge (e \vee \overline{\dot{e}}) & & \\
= & & \\
\overline{e} \wedge (e \vee \overline{\dot{e}}) & \geq & \overline{e} \wedge (e \vee \dot{e}) \\
 & & = \\
 & & \overline{e} \wedge \top \\
 & & = \\
 & & \overline{e}
\end{array}
$$

Page 44n

$$e$$
$$=$$
$$e+0$$
$$=$$
$$e+(1-e)0$$
$$=$$
$$e+(1-e)\overline{0}$$
$$=$$
$$e+(1-e)\overline{(1-e)e}$$
$$=$$
$$e+(1-e)\overline{\overline{(1-e)}\,\overline{e}}$$
$$=$$
$$e+(1-e)\overline{e}$$
$$=$$
$$e+\overline{e}-e\overline{e}$$
$$=$$
$$e+\overline{e}-e$$
$$=$$
$$\overline{e}$$

$$\lfloor x-\lfloor y\rfloor\rfloor \qquad\qquad \lfloor\lfloor x\rfloor-\lfloor y\rfloor\rfloor$$
$$0+\lfloor x-\lfloor y\rfloor\rfloor \qquad\qquad 0+\lfloor\lfloor x\rfloor-\lfloor y\rfloor\rfloor$$
$$-\lfloor y\rfloor+\lfloor y\rfloor+(0\vee(x-\lfloor y\rfloor)) \qquad\qquad -\lfloor y\rfloor+\lfloor y\rfloor+(0\vee(\lfloor x\rfloor-\lfloor y\rfloor))$$
$$-\lfloor y\rfloor+((\lfloor y\rfloor+0)\vee(\lfloor y\rfloor+(x-\lfloor y\rfloor))) \qquad\qquad -\lfloor y\rfloor+((\lfloor y\rfloor+0)\vee(\lfloor y\rfloor+(\lfloor x\rfloor-\lfloor y\rfloor)))$$
$$-\lfloor y\rfloor+\quad(\lfloor y\rfloor\quad\vee\quad x) \qquad\qquad -\lfloor y\rfloor+\quad(\lfloor y\rfloor\quad\vee\quad\lfloor x\rfloor)$$
$$-\lfloor y\rfloor+\quad(((y\vee 0)\vee x) \qquad = \qquad -\lfloor y\rfloor+\quad(((y\vee 0)\vee(x\vee 0))$$

$$\lfloor 0\rfloor$$
$$=$$
$$\lfloor\lfloor 0\rfloor\rfloor-\lfloor-\lfloor 0\rfloor\rfloor$$
$$=$$
$$\lfloor\lfloor 0\rfloor\rfloor-\lfloor 0-\lfloor 0\rfloor\rfloor$$
$$=$$
$$\lfloor\lfloor 0\rfloor\rfloor-\lfloor\lfloor 0\rfloor-\lfloor 0\rfloor\rfloor$$
$$=$$
$$\lfloor\lfloor 0\rfloor\rfloor-\lfloor 0\rfloor$$
$$=$$
$$\lfloor\lfloor\lfloor 0\rfloor\rfloor-\lfloor 0\rfloor\rfloor-\lfloor\lfloor 0\rfloor-\lfloor\lfloor 0\rfloor\rfloor\rfloor$$
$$=$$
$$\lfloor\lfloor 0\rfloor-\lfloor 0\rfloor\rfloor-\lfloor\lfloor\lfloor 0\rfloor\rfloor-\lfloor\lfloor 0\rfloor\rfloor\rfloor$$
$$=$$
$$\lfloor 0\rfloor-\lfloor 0\rfloor$$
$$=$$
$$0$$

Page 97n

$$\frown|(\frown|x)$$
$$(\odot|\top)|(\frown|x)$$
$$(\odot|\frown)|(\top|x)$$
$$((\top|\bot)|(\top|\odot))|(\top|x)$$
$$(\top|(\bot|\odot))|(\top|x)$$
$$(\top|\smile)|(\top|x)$$
$$\top|(\smile|x)$$

$$\frown|(\smile|x)$$
$$(\odot|\top)|(\smile|x)$$
$$(\odot|\smile)|(\top|x)$$
$$((\bot|\top)|(\bot|\odot))|(\top|x)$$
$$(\bot|(\top|\odot))|(\top|x)$$
$$(\bot|\frown)|(\top|x)$$
$$(\bot|\top)|(\frown|x)$$
$$\odot|(\frown|x)$$

$$(U-V)\cup(U'-V')$$
$$[(U-V)^C\cap(U'-V')] + [(U-V)\cap(U'-V')] + [(U-V)\cap(U'-V')^C)]$$
$$[(U^C\cap(U'-V')] + [V\cap(U'-V')] + [(U-V)\cap(U'-V')] + [(U-V)\cap U'^C] + [(U-V)\cap V']$$
$$[(X-U)\cap(U'-V')] + [(V-\emptyset)\cap(U'-V')] + [(U-V)\cap(U'-V')] + [(U-V)\cap(X-U')] + [(U-V)\cap(V'-\emptyset)]$$
$$[U'\backslash(U\cup V')] + [U'\cap V)\backslash V'] + [(U\cap U')\backslash(V\cup V')] + [U\backslash(U'\cup V)] + [(U\cup V')\backslash V]$$
$$[(U\cup U')-(U\cup V')] + [(U'\cap V)-(V\cap V')] + [(U\cap U')\backslash(V\cup V')] + [(U\cup U')-(U'\cup V)] + [(U\cup V')-(V\cap V')]$$

$$\int$$

Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104
`pjf@upenn.edu`

Available at
`http://www.math.upenn.edu/~pjf/analysis.pdf`
and check out
`http://www.math.upenn.edu/~pjf/e-pi.pdf`

Index

Department of Mathematics
University of Pennsylvania
Philadelphia, PA 19104
pjf@upenn.edu

Available at
http://www.math.upenn.edu/~pjf/analysis.pdf
and check out
http://www.math.upenn.edu/~pjf/e-pi.pdf