A Meaning Explanation for HoTT*

Dimitris Tsementzis

February 15, 2017

Abstract

The Univalent Foundations (UF) offer a new picture of the foundations of mathematics largely independent from set theory. In this paper I will focus on the question of whether Homotopy Type Theory (HoTT) (as a formalization of UF) can be justified intuitively as a theory of shapes in the same way that ZFC (as a formalization of set-theoretic foundations) can be justified intuitively as a theory of collections. I first clarify what I mean by an "intuitive justification" by distinguishing between formal and pre-formal "meaning explanations" in the vein of Martin-Löf. I then explain why Martin-Löf's original meaning explanation for type theory no longer applies to HoTT. Finally, I outline a pre-formal meaning explanation for HoTT based on spatial notions like "shape", "path", "point" etc. which in particular provides an intuitive justification of the axiom of univalence. I conclude by discussing the limitations and prospects of such a project.

Introduction

The Univalent Foundations (UF) is a new proposal for the foundations of mathematics that lies outside the spectrum of Cantorian foundations.[1] UF takes as its intended semantics the universe of ∞-groupoids (or homotopy types). It does so in terms of formal theories called Homotopy Type Theories (HoTT(s)). These are usually phrased in terms of some formal dependent type theory in the style of Martin-Löf. For example, the "standard" HoTT in [34] is given by intensional Martin-Löf Type Theory with some extra assumptions (the axiom of univalence and certain higher inductive types). However, it is important to note

*I would like to thank John Burgess, Hans Halvorson and Colin McLarty for helpful comments on earlier drafts. I would also like to single out in thanks Harvey Friedman, both for his many comments and questions, as well as for his penetrating criticism of the "new foundations" that led to a significant sharpening of the particular aims of this paper. Finally, I would like to thank two anonymous referees for highly engaged and illuminating remarks. This work was partially supported by NSF CAREER-DMS-1554092 (P.I. Harry Crane).

[1]For standard introductions to UF and HoTT, see [27, 30, 34, 39]. For the ideas that led to the homotopy interpretation of type theory see [5, 41]. For some of the earlier writing that led to the ideas for UF see [35–37]. For a discussion of the sense in which UF is a foundation of mathematics and connections to mathematical structuralism, see [4, 32, 33]. For some philosophical issues associated to HoTT see [11, 12, 19, 28].

that philosophically UF is not bound to dependent type theories even though dependent type theories have so far proven the best way to formalize UF.[2]

Set theory, formalized as ZFC, is the most widely-recognized foundation of mathematics and the mathematical community's accepted gold standard of rigour.[3] A pressing issue is therefore to articulate the relationship between set-theoretic foundations and the univalent foundations, as well as more specifically that of ZFC and HoTT (as their respective formalizations). There are two aspects to this relationship, a technical and a philosophical one. The technical relationship between UF and set theory has been made precise in various ways. For example, the simplicial model of univalence (cf. [18]) gives an interpretation of HoTT into ZFC supplemented with two inaccessible cardinals and Cubical Type Theory (CTT) (cf. [6, 10]) is expected to be interpretable into a constructive theory like CZF [1]. In the other direction, the 0-types ("*h*-sets") in a univalent universe are known to model (variants of) Lawvere's ETCS (cf. [20]), and there is also a way to construct a model of ZFC through a higher inductive type (cf. [34], 10.5).

The philosophical relationship is on the other hand less well-understood. This uncertainty leads to a rather obvious worry: if the basic objects of UF are best understood through their formalization in set theory then is UF really *autonomous* from set theory? And if it is not, then does it really deserve the name of a foundation? One way to make the question of autonomy precise is as follows: can the rules of HoTT (as a formalization of UF) be justified *pre-formally* in terms of notions that are independent from naive set theory?[4]

To answer this question requires picking some *specific* HoTT, which is then to be supplied with a pre-formal *meaning* in terms of which its rules can be *justified*. Carrying out this process is what I will understand by the phrase "a meaning explanation for HoTT". This is the task we undertake in this paper, using a meaning explanation that is *spatial* in character. This paper may therefore be seen as a realization of the suggestion made by Ladyman and Presnell [19] to justify HoTT by providing "an account of spaces that contains the features needed to support [the rules of HoTT], but which is grounded in pre-mathematical intuitions rather than homotopy theory."

The plan is as follows. In Section 1 I clarify what I mean by a meaning explanation and how it relates to other uses of the term. In Section 2 I explain why Martin-Löf's meaning explanation does not apply to HoTT and specify a specific HoTT to which the proposed meaning explanation will apply. Finally, in Section 3 I provide this meaning explanation in terms of primitive *spatial*

[2]Thus, even if we agree that ∞-groupoids are to be formalized in terms of some dependent type theory in the tradition of Martin-Löf, I will not, in general, take HoTT to refer to some fixed formal system (e.g. the one in [34]). Rather, I take HoTT to be a generic term referring to plausible (generally accepted) formalizations of the universe of ∞-groupoids. For example, Cubical Type Theory [6], Voevodsky's Homotopy Type System (HTS) and also the Angiuli, Harper, Wilson type theory in [3], are HoTTs in this sense.

[3]For an account of ZFC's emergence as a standard of rigour see [8].

[4]There are other, less specific, ways of understanding autonomy. The autonomy that I speak of is perhaps closest to the "justificatory autonomy" of foundations as made precise by Linnebo and Pettigrew [21].

notions, accompanying it with a description of the simplicial model of univalence on which it is based. I conclude with some discussion of the limitations and prospects of my approach in Section 4.

1 What is a Meaning Explanation?

The term *meaning explanation* is used in various often conflicting contexts and therefore requires clarification. A meaning explanation as I will understand the term applies to a *formal system* F. A *formal system* F is comprised of a formally specified syntax S and rules R of the form

$$\frac{s_1 \quad s_2 \quad \dots \quad s_n}{t} \; R$$

where $s_1, \dots, s_n, t$ are elements of S. For example, the s_i could be well-formed formulas if S is taken to be a first-order signature or they could be expressions of the form $\Gamma \vdash \mathcal{J}$ if S is the syntax of a dependent type theory.

A meaning explanation now does two things. Firstly, it assigns to each $s \in S$ a *meaning* s^M, i.e. some way of understanding s that goes beyond the symbols which comprise it. For example, for S a first-order signature and s the formula $\phi \wedge \psi$ then the classical meaning for s would be "ϕ and ψ are true". Or if S is a dependently typed syntax and s the expression $x\colon A \vdash B$ **Type** then a meaning for s would be that "in context $x\colon A$, B is a well-formed type".

Secondly, a meaning explanation *justifies* the rules R through their assigned meaning. This means that a rule R as above is justified if from the truth of $s_1^M, \dots, s_n^M$ the truth of t follows. For example, let S be a first-order signature, let R be the $\wedge$-intro rule

$$\frac{\phi \qquad \psi}{\phi \wedge \psi} \; \wedge\text{-intro}$$

and let the meaning of $\wedge$ be the usual conjunction. Then the rule R is justified by arguing that if ϕ is true and ψ is true then $\phi \wedge \psi$ is true. The same reasoning also applies to axioms, whose justification for the purposes of this paper should be taken to be a special case of the situation just described. For example, the formal system ZFC can be defined with S the set of well-formed first-order formulas over the signature with $\in$ the only non-logical symbol and with R the usual rules for classical first-order logic together with rules of the form

$$\frac{}{\phi} \; A$$

for every axiom ϕ of ZFC. A meaning explanation for ZFC would then need to give a meaning to the symbol $\in$ such that the axiom-rules such as A are justified, which in this case means simply that ϕ can be argued to be true (from no premises). For example, for ϕ the axiom for the empty set we would have to argue that it is true under the particular meaning of $\in$ that there is an entity with no elements. The same reasoning also applies to formal systems that do not fit into the mould of first-order logic, like the dependent type theories we will be concerned with for most of this paper. There are thus two key components of a meaning explanation for a formal system $F = (S, R)$: the *meaning* of the syntax S and *justification* of the rules R.

We must now distinguish between *formal* and *pre-formal* meaning explanations. A *formal* meaning explanation assigns to each s in S a meaning in terms of notions that are understood in terms of some other formal system. For example, a formal meaning for $\phi \wedge \psi$ would be given by interpreting ϕ, ψ as sets and $\wedge$ as their cartesian product and their respective truth as their inhabitation. What makes this a *formal* meaning is that in this case a *set* is understood as a formal construct in its own right, described by another formal system $F' = (S', R')$ (e.g. by ZFC) and so what we are really doing is interpreting F into F'. Similarly, interpreting the judgments of a dependent type theory as statements about a certain contextual category (cf. [9,38]) is to supply them with a *formal* meaning.

On the other hand, a *pre-formal* meaning explanation assigns to each s in S a meaning in terms of notions intuitively comprehensible even without a mathematical theory. For example, a pre-formal meaning for $\phi \wedge \psi$ would be given by interpreting $\wedge$ as the natural language "and". What makes this a *pre-formal* meaning is the fact that natural language is not to be regarded as another formal system F' but is rather a setting in which words have some *intuitive* content, comprehensible to the human mind. The difference between the formal and pre-formal approaches as I understand it is brought out very clearly by Martin-Löf in [25]:

> [Model Theory] and translation into another language are one and the same thing [...] [Therefore] you can give meaning in this way only if you have another language to translate into. [...] But, eventually, you will of course have to have a language which is not given meaning by translating it into another language but has to be given meaning in some other way, and this is the language of the most primitive notions that you are dealing with, because that they are primitive means precisely that they cannot be defined in terms of any other notions. ([25], p. 408)

Another way of explaining this difference is in terms of "pre-mathematical" versus "meta-mathematical" justifications for a formal system. A *pre*-mathematical justification is one that does not rely on an interpretation of the given formal system within some equally formal metatheory, whereas a *meta*-mathematical justification is one that interprets the given formal system in some other formal system. To use Dybjer's [13] example: a realizability model of MLTT is a meta-mathematical justification, whereas his own "testing semantics" constitute a pre-mathematical justification. The reader may take my use of the terms "intuitive" and "pre-formal" as synonymous to "pre-mathematical".

Whatever the relative merits of formal or pre-formal meaning explanations may be, it is very important for the purposes of this paper to keep them apart.

For what I am interested in here is a pre-formal meaning explanation for Homotopy Type Theory. I am in agreement with Martin-Löf that providing such a pre-formal meaning explanation is

> [a] genuinely semantical or meaning theoretical investigation, which means that [one] must enter on something that [one] is not prepared for as a mathematical logician, whether model theorist or proof theorist: [one] must enter on an enterprise which is essentially philosophical or phenomenological [...] in nature. ([25], p. 408)

Thus what I am after here is neither the kind of interpretation of HoTT in set theory achieved in [6, 18], nor the formalization of the "Initiality Conjecture" in terms of C-systems outlined in [38], nor even more recent efforts by Angiuli, Harper and Wilson [3] to provide computational meaning explanations to HoTT-style systems in terms of realizability semantics. What I am after, rather, is a genuinely "meaning-theoretical" explanation of HoTT. Therefore, the ultimate goal of this paper is to do for HoTT using notions such as "shapes", "points" etc. what Boolos [7] did for ZFC using notions such as "stages", "properties" etc.

2 What is a Meaning Explanation for HoTT?

HoTT is based on Martin-Löf Type Theory (MLTT).[5] MLTT differs from first-order logic in that the basic elements of its syntax ("judgments") are not to be understood as well-formed formulas lacking an interpretation, but are rather to be understood as interpreted statements *about* formulas. This difference is expressed nicely by Sundholm in [15], when discussing Martin-Löf's general philosophical outlook towards the foundations of mathematics:

> [MLTT] is an *interpreted formal language*. [...] [In first-order logic] the metamathematical "expressions" employed in mathematical logic are mere objects of study, but do not express. On the contrary, they are objects that may serve as referents of real expressions. In [MLTT], on the other hand, the expressions used are real expressions that carry meaning. In a nutshell, the language is endowed with meaning by turning the proof-theoretic reductions into steps of meaning explanation. Just like Frege's ideography, or of the language of *Principia Mathematica*, the type-theoretic formulae are actually intended to say something. They do not essentially serve as objects of metamathematical study. ([15], p. xx)

The judgements of MLTT, in other words, do not immediately serve as objects of metamathematical study because they are *already* metamathematical: they speak of the truth of propositions rather than being propositions themselves.

[5] This does not mean that any formalization for UF will be based on MLTT. I take up this point in the conclusion.

Thus, MLTT was not originally presented merely as a formal system with which to encode mathematics (although, of course, this was mainly what it was used to do) and this is why Martin-Löf also paid a lot of attention to the issue of what the basic syntax of MLTT *meant* and how its rules could be *justified*, i.e. how to give MLTT a pre-formal meaning explanation.

The original meaning explanation of Martin-Löf gave the basic elements of MLTT (the "judgments") a meaning in terms of notions of computation and evaluation to canonical values, mixed in with a BHK-style interpretation of the logical connectives and quantifiers.[6] For example, the meaning of judgment

$$A\ \mathbf{Type}$$

was that A is specification for a particular task or program and the meaning of

$$a\colon A$$

was that a is a program that satisfies the specification of A or, equivalently, which can carry out the task specified by A.[7] The rules of the system were then justified by arguing that they are valid under this way of understanding the judgments. In modern terms, the original meaning explanation of Martin-Löf can be understood as a pre-formal description of a realizability model for MLTT.

So why can we not build on Martin-Löf's meaning explanation in order to justify HoTT? Because, quite simply, the meaning explanation of Martin-Löf validates a rule for identity types ("Uniqueness of Identity Proofs" or UIP) with which HoTT is inconsistent. In particular, UIP is inconsistent with the axiom of univalence which is the fundamental new addition to HoTT that differentiates it from MLTT.[8] Thus, the grounds upon which the additional rules of HoTT are to be justified cannot be an extension of the grounds upon which the original rules were justified.

Therefore, the only option left to us is that of providing a completely new pre-formal meaning explanation for the rules of HoTT, one that neither piggybacks on set theory, nor extends Martin-Löf's programming-inspired meaning explanation. Furthermore, let me note that this also means that MLTT and HoTT should be viewed as fundamentally *distinct* formal systems and not species of the same genus (a genus one may broadly label "dependent type theories"). Although sociologically inevitable, to call HoTT a type theory is, I think, conceptually a mistake. Any formalization of UF should rather be called "∞-groupoid theory" in much the same way that ZFC (or CZF, or ZF etc.) is called a "set theory". Dependent Type Theories should strictly speaking be understood as those formal systems which can be given a Martin-Löf-style meaning explanation. And as I have argued above this is not true of any HoTT.

[6] For more details see [23, 26] as well as remarks in the original [24]. A good exposition can be found in [14].

[7] More precisely, that a has a *canonical* program of the canonical type A as a *value*.

[8] The justification for UIP goes roughly as follows: insofar as we have a proof of identity, we no longer care to distinguish between the two terms shown to be identical; therefore all that matters is *that* they are identical, which means that there should be only one canonical term demonstrating this fact. See [14], p. 8 for further discussion.

Since meaning explanations as I understand them apply to formal systems, in order to give a meaning explanation for HoTT we need to specify a suitable version of HoTT. We should pick one that is neither too weak nor too strong.[9] A reasonable choice of such a HoTT should in addition to the usual structural rules of MLTT contain at least three elements. Firstly, a univalent universe of types $\mathcal{U}$ because that is at minimum what distinguishes HoTT from MLTT. Secondly, we need the full expressive power of first-order logic and hence all the corresponding type constructors in MLTT ($\Pi, \Sigma, +, \mathtt{Id}, \mathbf{0}, \mathbf{1}$). Finally, we will clearly require the natural numbers $\mathbb{N}$, defined as an inductive type. Thus, a HoTT minimally sufficient for our purposes contains the following:

1. The usual syntax of MLTT consisting of the four forms of judgment.
2. The usual structural rules of MLTT governing contexts, substitution and judgmental equality.
3. Rules for Σ, Π, $\mathtt{Id}$, $+$, $\mathbf{0}$, $\mathbf{1}$ and $\mathbb{N}$.
4. A universe $\mathcal{U}$ closed under all the above type constructors satisfying univalence.

I will abbreviate this system as 1UHoTT (for "one universe HoTT"). So from now on "a meaning explanation for HoTT" is to be understood as a "meaning explanation for 1UHoTT."[10] Clearly this is not the only system that adequately formalizes UF, nor is it clearly the most suitable one to receive a meaning explanation. Indeed, as I will discuss below, Cubical Type Theory (CTT) might prove better-suited to this task. Furthermore, 1UHoTT is a rather weak system, since in the absence of W-types it is unable to encode many standard inductive definitions. But for our purposes here, this is not a direct concern: we are looking for the weakest possible system that can interpret arithmetic and be plausibly regarded as a HoTT. 1UHoTT fits this description well enough. A full list of rules for 1UHoTT is given in an Appendix.

3 A Spatial Meaning Explanation for 1UHoTT

The main idea the spatial meaning explanation for HoTT is to understand the expressions in the syntax as spatial observations. A little more precisely: the meaning of type constructors and term constructors consists in describing what can be done on previously constructed shapes and their points. For example, as we shall see, Σ-types are to be viewed as "gluing" shapes together and $\mathbb{N}$

[9] Exactly how to understand "strong" in the setting of dependent type theories is itself a complicated issue, cf. the proof-theoretic analyses in [29] and [16] for an illustration of the relevant techniques.

[10] UniMath [40] is the closest formal realization of 1UHoTT, at least insofar as one ignores the fact that it is built on top of Coq, which is itself based on the much stronger Calculus of Inductive Constructions. Alternatively, 1UHoTT can be taken as the system described in the Appendix of [18] if we replace the rules of W-types with the rules for $\mathbb{N}$ (which follow anyway as a special case of W-types).

is to be viewed as the infinite "discrete" shape. The rules of the system are then justified by arguing that moving from a given observation to a new one is correct based on spatial intuition.

So what is now required is, firstly, to use this idea to supply the basic syntax with pre-formal meaning and then, secondly, to use this meaning to justify the rules of 1UHoTT. However, given that there is also a formal meaning explanation of HoTT available in [18], and given that the outline of such a formal meaning explanation might be of independent interest to the reader and might help illuminate the accompanying pre-formal meaning explanation, I will in what follows accompany the pre-formal explanations with their corresponding formal ones, following the so-called *simplicial model* of HoTT constructed in [18]. I will do so by using notions that are as much as possible "foundation-agnostic" – "topological space", "fibration" etc. – but which nevertheless have precise counterparts in the simplicial model. So one can view the spatial meaning explanation that follows in two ways: either as an explanation of the simplicial model of univalence in naive terms, or as a pre-formal explanation together with a formal illustration in terms of a mathematical model. To be clear: the pre-formal meanings and justifications given below are supposed to be pre-formal analogues of the corresponding formal meanings and justifications and the latter are guided by the simplicial model of univalence. Therefore, one who believes in the simplicial model of univalence can simply understand what follows as an argument that it has intuitive content.

3.1 The Meaning of the Syntax

There are three basic elements of the syntax of MLTT whose meaning we need to explain: judgments, hypothetical judgments and substitution. I take them in turn.

3.1.1 Judgments

Judgments will be understood as *observations*.[11] There are four types of observations, corresponding to the four types of judgment in MLTT. They are the following, writing "**Shape**" instead of the usual "**Type**" to hint at the intended meaning:

1. $a : A$
2. $a = a' : A$
3. A **Shape**
4. $A = A'$ **Shape**

The meaning of the four kinds of observation is then given as follows:

[11] There is an unfortunate terminological clash with *Observational* Type Theory (OTT), a variant of MLTT considered in [2]. The terminological connection between my use of "observation" here and the use of the same word in OTT is a coincidence.

1. a is a **point** of A
2. a and a' are **symmetric** as points of A
3. A is a **shape**
4. A and A' are **symmetric** as shapes

There are thus three new basic notions, on which the whole proposed meaning explanation for 1UHoTT rests:

(a) **point** (replacing "term")

(b) **shape** (replacing "type")

(c) **symmetric** (replacing "judgmentally equal")

These three spatial notions (a)-(c) are I believe intuitive enough for a pre-formal meaning explanation. I will now explain what this intuition consists in. The key notion that is in need of explanation is that of a *shape* since what its *symmetries* and *points* mean will follow from such an explanation.

Pre-formal Meaning of Shape, Point, Symmetry: By a *shape* I understand a spatial entity built up *points* and *paths* between points. Paths and points are the only two primitive constituents of a shape. A shape of this kind is then understood to satisfy the following properties:

(Po) Shapes have *portions*. Such portions can be observed from points on other shapes and they are themselves shapes.

(PS) The paths between any two points on a shape also form a shape. We can call these *path shapes*. The points of path shapes are paths.

(Co) There are shapes that are indistinguishable from points. We can call these *contractible* shapes. A shape is contractible if and only if there is a point a on the shape such that from any point on the shape we can observe a path to a.

(C) For any given point a of a shape there is a shape consisting of all points b and paths from b to a. We can call these the *components* of a shape. Every component of a shape is contractible.

(PI) Two points in a shape are connected by a path if and only if they are indistinguishable with respect to any observation we can make from them.[12]

(MS) Between any two shapes A, B there is a shape of maps between them. A point of this map shape is a rule that assigns to each point of A a point of B. Such a rule can be obtained if for any point of A we can observe a point of B. We can call these *map shapes* and their points we can call *maps*.

[12] The biconditional here is crucial. It will be used essentially in the justification of `Id`-shapes.

(1) There exists a shape with a point.

(V) Between all distinct shapes there is a *void* which is understood as the absence of points. This entire void we may call the *empty shape*.

(N) There exists a shape with infinitely many points not connected by a path.

(MI) Any two shapes A and B are indistinguishable if and only if there is a map f from A to B that establishes a one-to-one correspondence between the points of B and the points of A. We can call such a map an *equivalence*.

(U) There is a shape that contains all shapes. We can call such a shape a *universe*. A path between two shapes (regarded as points of the universe) corresponds uniquely to an equivalence between them. Conversely, every equivalence between shapes uniquely determines a path between them in the universe.

The pre-formal meaning explanation below will rely only on the (intuitive content of) the above-listed properties of shapes. Properties (PS)-(PI) describe features of individual shapes. Properties (MS)-(U) describe features concerning the relations between shapes. Note also that the properties (MI) and (U) are only essential in order to justify the univalence axiom and (N) is only required to justify the rules for $\mathbb{N}$. So even if one finds them counterintuitive that still leaves a big chunk of rules that can be justified in terms of the rest of the properties. Furthermore, it must be noted that the notion of *indistinguishability* that I have employed above will be used as a primitive notion. It must be understood relative to the above-introduced notion of observation: two shapes or points are indistinguishable if there is no observation we can make from one that we cannot make from the other.

In broad strokes, a shape made up of points and paths and satisfying the above properties should be thought of as fully determined by the number of pieces it consists in and how each of these pieces relates one to the other. In other words, a shape can be thought of as a geometric object without a notion of distance or dimension but retaining information about how many pieces it consists in (formally, this is the "shapes-as-homotopy types" perspective). Shapes can be of any dimension (in the Euclidean sense) but their complexity is not at all captured by their dimension. Very high-dimensional objects (in the Euclidean sense) may be trivial as shapes and low-dimensional objects can be highly complex. As such, notions such as "convexity", "concavity" and "curvature" are irrelevant in this context. (They do not distinguish between shapes.) Visually, shapes are best thought of "higher graphs" i.e. as networks of nodes and edges and edges connecting edges etc. (formally, this is the "shapes-as-∞-groupoids" perspective).

A *point* is an irreducible component of a shape. One can think of these points as points in the Euclidean sense ("that which has no parts") except we do not necessarily regard them as the indivisible components of some ambient continuum, but rather as simply some primitive components. Importantly this means that one can observe a shape in this sense without observing any of its specific points, since it might not be clear from simply observing that something is a shape which of its components are indivisible. (This is what necessitates separate observations "A **Shape**" and "$a\colon A$").

Finally, by *symmetry* I understand the following: two shapes or points are symmetric if they are identical (*strict*, *Leibnizian* identity) but are being observed from a different point of view, where "observing them from a different point of view" here can be understood in terms of the symmetries of the shape itself, i.e. two distinct observations of the same shape can also be understood as an observation from the same point of view of a symmetric shape. (It is the difference between the "eye" moving in space, or the "eye" remaining constant while the shape itself flips.) For example, two points lying on opposite edges of a square, which coincide if we flip the square along the appropriate diagonal of symmetry.[13]

Formal Meaning of Shape: The formal meaning of a *shape* is a topological space up to homotopy equivalence.

A *point* of such a shape is then a point of a topological space.

A *symmetry* between two shapes is simply an equality (or even a homeomorphism) between them and a symmetry between two points is an equality.

Topological spaces up to homotopy have can be made precise in terms of the notion of ∞-groupoid. An ∞-groupoid can be thought of as a representation of a class of topological spaces that can be continuously deformed one into the other. Namely, an ∞-groupoid is what remains of a topological space however much we continuously deform it in such a way that doesn't change the number of pieces it consists in or the number of (possibly higher dimensional) holes it has. In set theory ∞-groupoids can be made precise in many ways. One such way which is used in [18] is as *Kan complexes*. These are defined within the category of simplicial sets, and are indeed the fibrant objects with respect to the standard Quillen model structure on simplicial sets. A *point* of a shape is then a point of a Kan complex and symmetry between shapes and points is given by equalities in the ambient category of simplicial sets. I will from now on call the formal versions of shapes *spaces*, with the understanding that this does not mean a general topological space, but rather a topological space up to homotopy equivalence.

Thus, to summarize, if the meaning of a judgment $\mathcal{J}$ in MLTT (under the original meaning explanation) is "to *know* that $\mathcal{J}$", the meaning of an observation in HoTT (under the spatial meaning explanation) is "to *observe* that $\mathcal{J}$", where now the $\mathcal{J}$ are interpreted in such a way that it makes intuitive sense to regard them as things one can observe (spatially). For example, instead of taking $a\colon A$ to mean that "we *know* that a is a program matching the specification

[13] It might be advisable to have two separate notions of symmetry, one for points (congruence) and one for shapes (symmetry). I have opted against this on grounds of parsimony – but the concerned reader should note that nothing I say below relies essentially on maintaining a single notion of symmetry, to be applied to points and shapes alike.

given by A" we take it to mean pre-formally that "we *observe* that a is a point of the shape A" and formally that "a is a point of a homotopy type".

3.1.2 Contexts and Hypothetical Observations

Hypothetical observations are expressions of traditional MLTT of the form $\Gamma \vdash x\colon A$ with Γ a *context*.

Pre-formal Meaning of Contexts and Hypothetical Judgments: Contexts will be understood as *viewpoints*. A viewpoint must be understood as what can be observed from an arbitrary point on a shape. Intuitively it tells us where exactly we can "move around" in order to make new observations. An expression $\Gamma \vdash \mathcal{O}$ for some observation $\mathcal{O}$ is to be understood as saying

"From viewpoint Γ, we can observe $\mathcal{O}$"

The empty context is to be understood as a point, indeed as the *basepoint* for any observation.

Expressions of the form $\Gamma \vdash x\colon A$ will therefore be understood as "expanded observations" with the intuition being that some observation $\mathcal{O}$ is made from a certain viewpoint Γ represented by the points and shapes on the left-hand side of the turnstile. Γ can be thought of as a shape which expands our capacity to make observations since we can move around it in order to observe other shapes/points/symmetries.

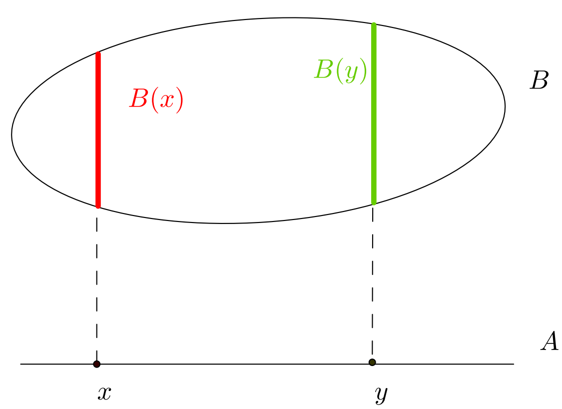
For example, consider the expression

$$x\colon A \vdash B(x) \textbf{ Shape}$$

The meaning of this expression is to be understood as follows:

"From any point x of A, we can observe a portion $B(x)$ of the shape B"

The fact that we have such portions and that they are shapes is guaranteed by (Po). We can visualize this idea as follows:



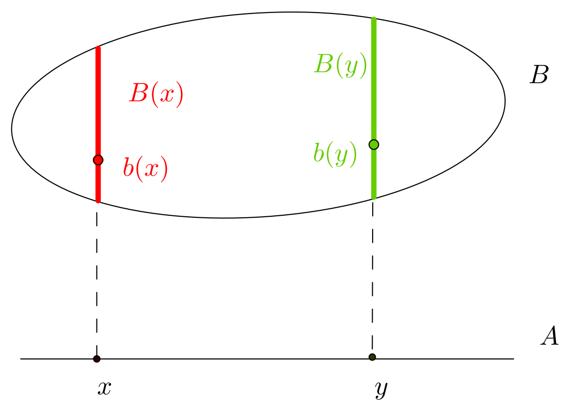
Similarly, consider the expression

$$x\colon A \vdash b(x)\colon B(x)$$

The meaning of this expression is to be understood as follows:

"From any point x of A, we can observe a point $b(x)$ of $B(x)$."

Once again, by (Po) we know that these portions are indeed shapes and that therefore it makes sense to speak of their points. We can visualize this idea as follows:



Thus, the meaning of type dependency in our setting becomes the following: a type B dependent on another type A is to be understood as a shape patched together from the different portions $B(x)$ that we can observe through arbitrary points x on A. As hopefully suggested by the above pictures, the mental image here should be one where from an arbitrary point x on the line A we shine a beam on the ellipse B and $B(x)$ is the line that we can now observe on the surface of B. However, it should be made clear that by "shining a beam" we do not all at once observe all the points of the portion of the shape that becomes visible to us. We merely observe just that: a portion (or component) of the whole shape.

It is important to note also that the "lines" $B(x)$ and $B(y)$ are in a sense symmetric. But we should certainly not take that to mean that we can assert the observation $B(x) = B(y)$ **Shape**. (Being able to do so would destroy the distinctive features of HoTT.) A symmetry between two shapes, recall, is understood as two distinct observations of the *same* shape. A symmetry, thus, is also a strict identity, and here $B(x)$ and $B(y)$ are not strictly identical.[14]

Formal Meaning of Contexts and Hypothetical Judgments: The formal meaning of a judgment $x\colon A \vdash B(x)$ **Shape** is that B is a space fibered

[14] I am indebted to an anonymous referee for clarifying my thinking on this point.

over A. In classical topology this is formalized as a *fibration* $\pi\colon B \to A$ whose *fibers* are given by $\pi^{-1}(x)$ for each $x \in A$.

The meaning of a hypothetical judgment $x\colon A \vdash b(x)\colon B(x)$ is that b is a *section* of the fibration $\pi\colon A \to B$, i.e. a map $s\colon B \to A$ such that $\pi \circ s = 1_A$. The reason why this gives a good notion of a "point" of B is because a section can be thought of as a rule for picking a point over the fiber of x in A for any x in a "continuous" way.

3.1.3 Substitution

Pre-formal Meaning of Substitution: Given a certain point a on a shape A, and any expression B containing a variable x of type A, the expression $B[a/x]$ is to be understood as the expression (whether shape or point) in which we have fixed the specific point a in place of arbitrary points x on A. Substituting specific points inside an expression is thus to be understood as "fixing" or "constricting" a viewpoint down to a single point.

Formal Meaning of Susbtitution: Substitution is understood as pullback of a fibration along a section. In more detail, assume we have $\Gamma \vdash a\colon A$ and $\Gamma, x\colon A \vdash b(x)\colon B$. This means that we have a section $a\colon \Gamma \to A$ and also a section $b\colon A \to B$. We thus obtain a pullback diagram

$$\begin{array}{ccc} \Gamma.B & \longrightarrow & B \\ \big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ \Gamma & \xrightarrow{\;a\;} & A \end{array}$$

and therefore a fibration $\Gamma.B \to \Gamma$ which we take as the interpretation of the observation $\Gamma \vdash B[a/x]$ **Shape**. An analogous construction using pullbacks gives us the interpretation of $\Gamma \vdash b[a/x]\colon B[a/x]$ (which has to be understood as a section $\Gamma \to \Gamma.B$). However, with this definition of substitution we now face the so-called "coherence problem" which arises from the fact that if substitution is interpreted as pullback in the manner just shown then successive substitutions do not commute strictly, but only up to isomorphism, contradicting the strictness required by the syntax. In other words, with substitutions-as-pullback we are not guaranteed to get the same thing if we substitute a for x and then b for y or if we substitute a and b for x and y simultaneously – we are only guaranteed to get two things that are isomorphic to each other. In order to solve this we must introduce a "strictification" of the interpretation, and this can be achieved in many ways. A construction using universes is how this is achieved in the setting of simplicial sets, but it is far beyond the scope of this paper to go into the details.

3.2 The Justification of the Rules

Given the interpretation of judgements as observations, of hypothetical judgements as expanded observations and of substitution as fixing a viewpoint in

Section 3, I will now move to the justification the rules of 1UHoTT. In some cases (Π, Σ and `Id`-shapes) I will first state fully and then give the justification of the rules. In other (less difficult) cases I will refer to the rules by name without stating them and give brief descriptions of how they are justified. The reader curious to see these rules stated (in order to perhaps convince themselves that they can be justified in the manner suggested) is referred to the Appendix.

3.2.1 Structural Rules

Pre-formal Justification: The rules for judgmental equality are justified by the fact that symmetry is clearly intuitively an equivalence relation, and that it "respects typing" because if a shape can be placed one on top of the other, then any point on either shape can also be seen as a point on the other shape. For example, take the following rule

$$\frac{\Gamma \vdash a\colon A \qquad \Gamma \vdash A = B \ \textbf{Shape}}{\Gamma \vdash a\colon B} \quad \texttt{Sym-tran}$$

What we have above the line is the assurance that we can observe a point a on a shape A and that shape A is symmetric to shape B. Being symmetric means that these two shapes are the same shape but just observed from a different point of view – think of A as a square and B as the square produced by flipping A along a line of symmetry. Then, clearly, the point a on A can also be observed to lie on B, thus giving us the observation below the line, and thus justifying `Sym-tran`. The justification of the rest of the structural rules concerning symmetry proceed very similarly.

The justification of `Vble` is immediate: if we can observe a point, then we can observe the same point. For $\texttt{Subst}_1$ we need to justify the following rule:

$$\frac{\Gamma \vdash a\colon A \qquad \Gamma, x\colon A, \Delta \vdash b\colon B}{\Gamma, \Delta[a/x] \vdash b[a/x]\colon B[a/x]} \quad \texttt{Subst}_1$$

What we have above the line is the assurance that we can observe a point a on a shape A and that from a viewpoint that includes an arbitrary point on A we can observe a point b on B. Below the line we have the assurance that we can observe a point $b[a/x]$ on $B[a/x]$ from the same viewpoint as above, except instead of an arbitrary point on A we now use the specific point a that we already know we can observe. Clearly, if we fix a viewpoint on a specific point, then every observation we could make from an arbitrary point we can also make from that specific point too, thus justifying the rule in question. The justification for $\texttt{Subst}_2$ proceeds similarly.

For $\texttt{Wkg}_1$ we need to justify the following rule:

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, \Delta \vdash b \colon B}{\Gamma, x \colon A, \Delta \vdash b \colon B} \ \texttt{Wkg}_1$$

What we have above the line is the assurance that we can observe a shape A from viewpoint Γ and a point b from viewpoint Γ, Δ. Now since the point b is observable from Γ, Δ it is certainly also observable from $\Gamma, x \colon A, \Delta$ since, on the assumption that this latter context is well-formed, the observation of b will proceed just as before. In other words, we already know what we need to observe b, and as long as an arbitrary point of A is observable from Γ, the rest of the points we need to observe b are already there, and the addition of x doesn't change the situation at all. The justification of $\texttt{Wkg}_2$ proceeds similarly.

Formal Justification: The formal justification for all the structural rules not involving substitution is immediate from the fact that symmetry is interpreted as the actual equality of points and of topological spaces, both of which are clearly equivalence relations. The rules involving substitution are also justified but only if we have some canonical way of picking *specific representatives* for the pullbacks in terms of which substitution is defined. This indeed can be done in the cases of interest for the interpretation of type theory, and in particular can be done in the setting of simplicial sets.

It should be noted that we also have structural rules for each specific shape constructor (Π, Σ, etc.) asserting that the conclusions of the rules are invariant under judgmental equality. These can be straightforwardly justified (both formally and pre-formally) assuming the justification of the rules of each specific such type constructor, given below.

3.2.2 Rules for Shape Constructors

The shape constructors of 1UHoTT are the following: Σ, Π, $\texttt{Id}$, $+$, $\mathbf{0}$, $\mathbf{1}$, $\mathcal{U}$, $\mathbb{N}$. The four types of rules (Formation, Introduction, Elimination, Computation) must now be understood as describing shapes. I see them as doing this by answering the following questions:

- **Formation**: When can we observe A?
- **Introduction**: When can we observe a point of A?
- **Elimination**: When can we observe points on another shape from A?
- **Computation**: What are the symmetries of A?[15]

[15] Uniqueness principles (e.g. for Π-types) can also be seen to answer the same question as Computation rules, and this is how I will regard them.

Note that the formation rule for a new shape constructor X can only be justified by providing the *meaning* of X. This is inevitable, since we cannot get anything for free and X is really a new piece of syntax that is irreducible to the syntax we had previously. Thus, in what follows, the justification of the formation rule will also include an explanation of the meaning of the type constructor X (both pre-formal and formal).

The `Id`-shapes are an exception to this rule since the notions required to explain their meaning have already been introduced in the explanation of the meaning of a shape, e.g. in (PS). This is reasonable because `Id`-shapes describe the "internal" structure of a shape and therefore determine the very meaning of a shape. Shape constructors such as Π and Σ on the other hand determine operations that can be carried out on shapes and therefore assume that the meaning of these shapes is already determined. This fundamental distinction is not very well brought out in the syntax of type theory, but becomes clear in the attempted meaning explanation.

Finally, in what follows I will generally forgo mention of the (always assumed) implicit viewpoint Γ. The reader may assume that all these justifications take place relative to a given viewpoint and I do not believe this alters their essential content.

Σ-shapes

$$\frac{\Gamma \vdash A \textbf{ Shape} \qquad \Gamma, x\colon A \vdash B(x) \textbf{ Shape}}{\Gamma \vdash \sum_{x\colon A} B(x) \textbf{ Shape}} \ \Sigma\texttt{-Form}$$

Pre-formal Justification: We assume that we can observe a space A and from each point x of A we can observe a shape $B(x)$ that is a portion of a shape B. We then take the shape $\sum_{x\colon A} B(x)$ to be formed by the operation of *attaching* (or *juxtaposing*) each shape $B(x)$ to the point x. And the rule is justified by appealing to our intuition of the meaning of *attaching* (or *juxtaposing*) as an operation that produces a shape given that A and each $B(x)$ are shapes.

Formal Justification: Given a space A and a fibration $\pi\colon B \to A$ we can always form the *total space* of the fibration and this is exactly what we take as the interpretation of $\sum_{x\colon A} B(x)$. And indeed the total space of a fibration can profitably be thought of as the space obtained by *attaching* to each point a of the base A the fiber $\pi^{-1}(a)$ in B over a.

$$\frac{\Gamma \vdash A \textbf{ Shape} \qquad \Gamma, x\colon A \vdash B(x) \textbf{ Shape}}{\Gamma, x\colon A, y\colon B(x) \vdash \langle x, y\rangle \colon \sum_{x\colon A} B(x)} \ \Sigma\texttt{-Intro}$$

Pre-formal Justification: We assume that we can observe a space A and from each point x of A we can observe a shape $B(x)$ that is a portion of a shape

B. Now, assume that we can observe a point x on A and a point y on $B(x)$. Then we can clearly observe a point on the shape $\sum_{x:A} B(x)$ since it is obtained by attaching each $B(x)$ on to x. In other words, to observe a point of $\sum_{x:A} B(x)$ is to observe a point y on $B(x)$ but in order to observe a point on $B(x)$ we also need to observe x.

Formal Justification: The total space of a fibration $\pi\colon B \to A$ has as points pairs $\langle a,b \rangle$ where a is a point of A and b is a point of the fiber $\pi^{-1}(a)$. Thus, given any two such points x in A and y in $B(x)$ we can observe a point $\langle x,y \rangle$ of the total space, and this is exactly what the bottom line expresses.

$$\frac{\begin{array}{c}\Gamma, z\colon \sum_{x:A} B(x) \vdash C \ \textbf{Shape} \\ \Gamma, x\colon A, y\colon B(x) \vdash d(x,y)\colon C(\langle x,y \rangle)\end{array}}{\Gamma, z\colon \sum_{x:A} B(x) \vdash \texttt{split}_d(z)\colon C(z)} \quad \Sigma\texttt{-Elim}$$

Pre-formal Justification: We assume that we can observe a portion $C(z)$ of a shape C from any point z of $\sum_{x:A} B(x)$ and also that we can observe a point $d(x,y)$ on $C(\langle x,y \rangle)$ (the portion of C observable from the point $\langle x,y \rangle$) for any x on A and y on $B(x)$. Now let z be any point of $\sum_{x:A} B(x)$ that we can observe. The very fact that we can observe z means that we can observe a point y on some $B(x)$ and therefore also x. But then, by assumption, we can also observe the point $d(x,y)$ on $C(\langle x,y \rangle)$ and since z is fully determined by x and y this also means that we observe a point on $C(z)$ which we call $\texttt{split}_d(z)$.

Formal Justification: Let C be any space fibered over the total space of a fibration, i.e we are given a fibration $\pi\colon C \to \sum_{x:A} B(x)$. Then by the universal property of the total space of a fibration in order to produce a section of π it suffices to produce a map d that takes any x in A and y in $B(x)$ to a point $d(x,y)$ in the fiber of C over the pair $\langle x,y \rangle$. And this is exactly what the rule expresses.

$$\frac{\begin{array}{c}\Gamma, z\colon \sum_{x:A} B(x) \vdash C \ \textbf{Shape} \\ \Gamma, x\colon A, y\colon B(x) \vdash d(x,y)\colon C(\langle x,y \rangle)\end{array}}{\Gamma, x\colon A, y\colon B(x) \vdash \texttt{split}_d(\langle x,y \rangle) \equiv d(x,y)\colon C(\langle x,y \rangle)} \quad \Sigma\texttt{-Comp}$$

Pre-formal Justification: This rule expresses exactly the fact "z is fully determined by x and y" which we invoked in the justification of Σ-`Elim`.

Formal Justification: This follows from the uniqueness of the map produced by the universal property of the total space.

Π-shapes

$$\frac{\Gamma \vdash A\ \textbf{Shape} \qquad \Gamma, x\colon A \vdash B(x)\ \textbf{Shape}}{\Gamma \vdash \prod_{x\colon A} B(x)\ \textbf{Shape}}\ \texttt{Π-Form}$$

Pre-formal Justification: We can observe a shape A and from any x on A we observe a portion $B(x)$ of a shape B. By (MS) we know that between any two shapes there is a shape of maps between them. Since portions of a shape are also shapes, we can take $\prod_{x\colon A} B(x)$ to be the shape of maps from A to the portions of B visible from each point x of A.

Formal Justification: Given any fibration $\pi\colon B \to A$ we have the space of its sections, i.e. the space of maps $s\colon A \to B$ such that $\pi s = 1_A$. This space of sections is exactly what we take as $\prod_{x\colon A} B(x)$.

$$\frac{\Gamma, x\colon A \vdash b(x)\colon B(x)}{\Gamma \vdash \lambda(x\colon A).b\colon \prod_{x\colon A} B(x)}\ \texttt{Π-Intro}$$

Pre-formal Justification: We assume that for any point x on A we can observe a point $b(x)$ on the portion $B(x)$. But by (MS) a map between shapes is nothing other than a rule that takes a point of one shape to a point of the other shape. Therefore we have a map from each point x on A to B such that the point x gets sent to the portion $B(x)$. But such a map is exactly a point of $\prod_{x\colon A} B(x)$ as just described.

Formal Justification: If we are given a rule that produces a point of $B(x)$ for every x in A then this gives us a section of the fibration $B \to A$. This is exactly what the rule expresses.

$$\frac{\Gamma \vdash f\colon \prod_{x\colon A} B(x) \qquad \Gamma \vdash a\colon A}{\Gamma \vdash f(a)\colon B[a/x]}\ \texttt{Π-Elim}$$

Pre-formal Justification: We assume we can observe a point f on $\prod_{x\colon A} B(x)$ and also a point a on A. But since f is a map we know there we can observe a point on the portion $B(a)$ of B visible from a.

Formal Justification: Given a section f of a fibration $B \to A$ and a point a of the base space A then we can produce a point $f(a)$ in the fiber $B(a)$ by function application. This is exactly what the rule expresses.

$$\frac{\Gamma, x\colon A \vdash b\colon B(x) \qquad \Gamma \vdash a\colon A}{\Gamma \vdash (\lambda(x\colon A).b)(a) = b[a/x]\colon B[a/x]} \quad \Pi\texttt{-Comp}$$

Pre-formal Justification: We assume that we can observe a point $b(x)$ from any point x of A and that we are given a point a of A. From the first assumption we can, on the one hand, produce a point of $\prod_{x\colon A} B(x)$ and then this point will give us a corresponding point on $B(a)$. On the other hand, from the first assumption we know that from a some point will be observable on $B(a)$. These two points are clearly the same point, and this is what the rule expresses.

Formal Justification: Applying a to the section which takes x to $b(x)$ results in $b(a)$. This is exactly what the rule expresses.

Id-shapes

$$\frac{\Gamma \vdash A \textbf{ Shape} \qquad \Gamma \vdash a\colon A \qquad \Gamma \vdash b\colon A}{\Gamma \vdash \texttt{Id}_A(a,b) \textbf{ Shape}} \quad \texttt{Id-Form}$$

Pre-formal Justification: We assume that we can observe a shape A and points a and b on A. From (PS) we know that given any two points of a shape we can observe the shape of *paths* between these two points. This *path shape* is exactly what we take as the meaning of $\texttt{Id}_A(a,b)$.

Formal Justification: For any two points a, b of a space A we can form the path space $\textbf{Path}_A(a,b)$. In classical topology this path space can be defined as the set of continuous maps $p\colon [0,1] \to A$ with a and b as endpoints (i.e. such that $p(0) = 1$ and $p(1) = b$). The homotopy type of that path space is then the intended interpretation of $\texttt{Id}_A(a,b)$.

$$\frac{\Gamma \vdash A \textbf{ Shape} \qquad \Gamma \vdash a\colon A}{\Gamma \vdash \texttt{refl}_a\colon \texttt{Id}_A(a,a)} \quad \texttt{Id-Intro}$$

Pre-formal Justification: We assume that we can observe a shape A and a point a on A. A point is indistinguishable from itself and therefore any observation we can make from a can also be made from a. Therefore, by (PI) we get a path from a to a and therefore a point of the path shape which we call $\texttt{refl}_a$ and refer to as the *trivial path*.

Formal Justification: For any point a of a space A there is a trivial path from a to a that remains constant on a. Namely, the path $p\colon [0,1] \to A$ such that $p(x) = a$ for all $x \in [0,1]$.

$$\frac{\Gamma, x\colon A, y\colon A, p\colon \mathtt{Id}_A(x,y) \vdash C(x,y,p)\ \textbf{Shape} \quad \Gamma, z\colon A \vdash d(z)\colon C(z,z,\mathtt{refl}_z)}{\Gamma, x\colon A, y\colon A, p\colon \mathtt{Id}_A(x,y) \vdash J_{z.d}(x,y,p)\colon C(x,y,p)}\ \mathtt{Id\text{-}Elim}$$

Pre-formal Justification: We assume that from any path p on the path shape of x and y we can observe a portion $C(x,y,p)$ of a shape C and that from any point z on A we can observe a point $d(z)$ on the portion of C visible from the trivial path. Now let us assume that we can observe a path p in the path shape of x and y. By (C) this means we can observe a point $\langle y,p\rangle$ of the component of x. But we already know that from the component of x we can observe a point on the portion of C visible from the trivial path on x. By (PI) we know that x and y must be indistinguishable with respect to the observations we can make from them and therefore we should be able to observe a point on the portion of C visible from $\langle y,p\rangle$. We call this point $J_{z.d}(x,y,p)$.

Formal Justification: The formal justification of the `Id-Elim` rule was one of the key insights that led to the development of the field of homotopy type theory, due independently to Awodey-Warren [5] and Voevodsky. The main idea is that the factorization system in any Quillen model category gives us the necessary data to interpret identity types by taking the identity type to be the object given by the factorization of the diagonal map $\Delta\colon A \to A \times A$. In particular this structure exists given the model structure of the category of simplicial sets. For introductions to this idea, and detailed explanations of how it works, cf. [5, 18, 27].

$$\frac{\Gamma, x\colon A, y\colon A, p\colon \mathtt{Id}_A(x,y) \vdash C\ \textbf{Shape} \quad \Gamma, z\colon A \vdash d(z)\colon C(z,z,\mathtt{refl}_z)}{\Gamma, x\colon A \vdash J_{z.d}(x,x,\mathtt{refl}_x) = d(x)\colon C(x,x,\mathtt{refl}_x)}\ \mathtt{Id\text{-}Comp}$$

Pre-formal Justification: We assume that from any path p on the path shape of x and y we can observe a portion $C(x,y,p)$ of a shape C and that from any point z on A we can observe a point $d(z)$ on the portion of C visible from the trivial path. The rule now says that the point $d(z)$ is congruent to the point we can observe by the general rule described in the justification of `Id-elim` when that rule is applied to the trivial path. These two points are clearly the same points, and this is what the rule expresses.

Formal Justification: The rule is justified by the uniqueness of factorizations in the corresponding factorization system in the category of simplicial sets.

+, 0 and 1-shapes

Pre-formal Justification: For + (i.e. coproduct types) we use the idea of "concatenation" and the rules are justified using just this intuition, taking care to note that "concatenation" here is not left-right symmetric, i.e. that it matters which shape is concatenated on the "left" and which one on the "right", as of course one would expect from the usual behaviour of coproducts which separate points depending on which component they originate from. We regard $\mathbf{0}$ as an empty (or point-less) shape as described in (V). We regard $\mathbf{1}$ as a shape that consists of a single point. Its existence follows from (1) and (C), since if there exists a shape with a point then the component of that point is a shape indistinguishable from a point, and that shape is exactly what we take as the meaning of $\mathbf{1}$.

As an illustration of the justification of the rules we do what is arguably the most difficult case, namely that of `0-Elim`, which is given as follows:

$$\frac{\Gamma, x\colon \mathbf{0} \vdash C(x)\ \textbf{Shape}}{\Gamma, x\colon \mathbf{0} \vdash \texttt{efq}\colon C(x)}\ \texttt{0-Elim}$$

We assume that from a point on the empty shape we can observe a portion of a shape C. But a point x of the empty shape would be a point "in the void" between shapes, as explained in (V). And so if from that point you can observe a portion of a shape then there is nothing separating you from "moving closer" to that portion in order to observe a point (since you can freely move in the void, having been granted *per impossibile* a point on the void). The intuition – at the risk of sounding fanciful – is that if you are given a point of the empty shape (the "void") then you are given a foothold to make observations that you normally cannot make since there is nowhere to "stand on", so to speak. Therefore, given a place to "stand on" in the void you can move around the void freely, making every observation possible. But, of course, you should never be able to stand on a point in the void, since the void is the absence of points. In other words, the spatial interpretation of *ex falso quodlibet* is that having a point in the void allows you to observe anything whatsoever.

Formal Justification: We take $\mathbf{1}$ to be the singleton space, $\mathbf{0}$ to be the empty space, and the coproduct of two spaces to be given by their coproduct as spaces in the category of simplicial sets. The rules can then easily be seen to express the usual universal properties for each of these constructions in the category of simplicial sets.

$\mathbb{N}$-shapes

Pre-formal Justification: $\mathbb{N}$ is to be understood as an infinite discrete shape, i.e. a shape where there are infinitely many points not connected by paths. Spatially, this is not hard to imagine. But the main difficulty with justifying the rules for $\mathbb{N}$ is, roughly, that we need to sneak in some kind of "temporal" notion in our justification. In doing so, our justification ceases to be purely "shape-theoretic". This temporal element has to do with the need to talk about an "indefinitely extensible" process in order to assert the existence of arbitrarily many points in $\mathbb{N}$. Formally, this comes up in the introduction rule for $\mathbb{N}$ that involves the successor:

$$\frac{\Gamma \vdash n \colon \mathbb{N}}{\Gamma \vdash s(n) \colon \mathbb{N}} \quad \mathbb{N}\text{-}\mathtt{Intro}\text{-}2$$

Spatially, the rule $\mathbb{N}$-`Intro`-2 says that if we can observe a point n on $\mathbb{N}$ then we can observe another point $s(n)$. In other words, we need to somehow be able to say that the shape $\mathbb{N}$ is "indefinitely extendible".[16] Of course, indefinite extensibility can be cashed out geometrically in terms of the indefinite *divisibility* of some kind of continuum. But such an approach seems hopeless in our setting, since it requires us to assume an intuition of the continuum in order to describe a discrete shape such as $\mathbb{N}$. This might be a bullet worth biting, but only if other options have been exhausted. Therefore, I am currently not sure how to describe an infinite discrete shape (such as $\mathbb{N}$ is supposed to be) without either cashing out indefinite extensibility in terms of some temporal notion of progression, or relying on some ambient continuum which is indefinitely divisible. It seems to me that spatial intuition about shapes of the kind we have been talking about does not guarantee the existence of an infinite discrete shape that would justify the rules for $\mathbb{N}$. At the very least this situation seems no worse than the need to postulate an axiom of infinity in set theory.

Formal Justification: $\mathbb{N}$ can be constructed as a *natural numbers object* (NNO) in the category of simplicial sets. Such a NNO exists in this category because it is inherited from the category of sets. More generally, all W-types can be modelled in the category of simplicial sets (and any appropriately structured contextual category) as initial algebras for appropriate polynomial endofunctors. $\mathbb{N}$ is a specific instance of such an initial algebra.

$\mathcal{U}$-shape

Pre-formal Justification: A universe must be understood as some kind of "shape of shapes" and its existence follows from (U). However, intuitively, such a "shape of shapes" seems more difficult to conceptualize than a "collection of collections" which is how a universe in traditional MLTT can be understood. I will offer here a few suggestions as to how to understand the rules for the universe $\mathcal{U}$ as describing such a shape of shapes. The justification of the rules expressing the closure of the universe under specific type constructors can be justified just as before, so I will say no more about that. The other two rules are the following:

[16] Essentially, the task of making this precise amounts to giving an interpretation of PA/HA into a theory with pure geometric content. In the setting of first-order logic, this is not entirely without precedent. See for example the work of Hellman and Shapiro [17]. More recently, in private communication, Harvey Friedman has suggested a way of interpreting Z_2 (roughly, second-order arithmetic) into a first-order theory built out of purely geometric notions. Such efforts, especially the ones of Friedman that were motivated by similar considerations as the ones that motivate me here, could certainly help in fleshing out a purely geometric understanding of PA/HA, one that could then be used to give a more purely spatial justification to rules such as $\mathbb{N}$-`Intro`-2.

$$\frac{}{\vdash \mathcal{U}\ \textbf{Shape}}\ \texttt{Uni-Form} \qquad \frac{}{x\colon \mathcal{U} \vdash \texttt{El}(x)\ \textbf{Shape}}\ \texttt{Uni-El}$$

The rule `Uni-Form` is to be understood as asserting that a shape of shapes is observable from the "neutral" viewpoint, which is to say from *every* viewpoint. The visual that I find helpful here is that of a fractal-like accumulation of points – and the closer one "zooms into" one of these points, the more one sees them as shapes in their own right. Of course, this requires that a "point" is no longer necessarily understood as an irreducible part of a shape, but as a shape in its own right. And indeed, this is exactly what the dependent shape `El` expresses in `Uni-El` and this is also the reason why it is necessary to introduce it, viz. in order to be able to "blow up" points into shapes. Indeed, in line with (U), `Uni-El` asserts that from any point in the universe we can observe a shape, and we will take that shape to be the that very point from which it is observe. In other words: every point in a universe can regard itself as a shape. Admittedly, the problem with this talk of "fractal-like" self-similarity and of points morphing into shapes is that these concepts stray too far from what might be considered "intuitive", at least in the sense in which I have so far been using the term. I am currently not sure if a more intuitive visualization can be found.

Formal Justification: Given inaccessible cardinals $\alpha < \beta$ we obtain an appropriate notion of α-small Kan complexes and these can be shown to form a universe $\mathcal{U}_\alpha$ closed under all the shape-constructors and thus validating all the rules for a universe. This is recorded, for example, as Theorem 2.3.4 in [18].

Univalence for $\mathcal{U}$

$$\frac{\Gamma \vdash A\colon \mathcal{U} \qquad \Gamma \vdash B\colon \mathcal{U}}{\Gamma \vdash u_{A,B}\colon \texttt{isequiv}(\texttt{idtoequiv}_{A,B})}\ \texttt{Univalence}$$

Before we even begin the justification of the axiom of univalence, we need to explain the terms "`isequiv`" and "`idtoequiv`" in terms of which it is phrased. So, in the notation of HoTT we have:

$$\texttt{isequiv}(f) =_{\mathrm{df}} \prod_{x\colon A} \texttt{iscontr}(\texttt{fib}(f,x))$$

where for any type A we define

$$\texttt{iscontr}(\texttt{A}) =_{\mathrm{df}} \sum_{a\colon A} \prod_{x\colon A} \texttt{Id}_A(x,a)$$

and for $f\colon A \to B$ and $y\colon B$ we define

$$\texttt{fib}(f,y) =_{\mathrm{df}} \sum_{x\colon A} \texttt{Id}_B(f(x),y)$$

Also, setting

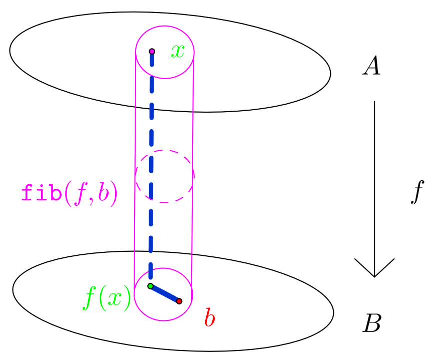
$$A \simeq B \equiv \sum_{f:\, A \to B} \texttt{isequiv}(f)$$

we have that

$$\texttt{idtoequiv}_{A,B} \colon \texttt{Id}_{\mathcal{U}}(A,B) \to A \simeq B$$

is the canonical term obtained by induction on identity by a standard argument.[17]

Pre-formal Justification: First we need to argue that the syntactic definition of an equivalence above corresponds to the intuitive definition in (MI), namely as a map that establishes a one-to-one correspondence between points of a shape. To see this, note that the fiber of a map f from A to B over a point b of B can be understood as a "cylinder shape" connecting those points x of A such that under f there is a path from $f(x)$ to b. This can be visualized as follows:



The cylinder traced out in magenta can be understood as the fiber of f over b, i.e. as the visual counterpart of the shape that the formal description of $\texttt{fib}(f,b)$ above describes. A point of this fiber can be visualized as the pair of blue lines above, i.e. a line (not a path!) connecting x to $f(x)$ together with a path from $f(x)$ to b. So to say that this fiber is contractible is to say that all these points can be squashed down to a single point. Since the points of the fiber can be thought of as a line attached to a path, then "squashing down to a point" for the fiber would correspond to being able to contract the cylinder down to a single line, connecting two points:

[17] First, we can regard $\simeq$ as a dependent type over any identity type in $\mathcal{U}$ by noting that

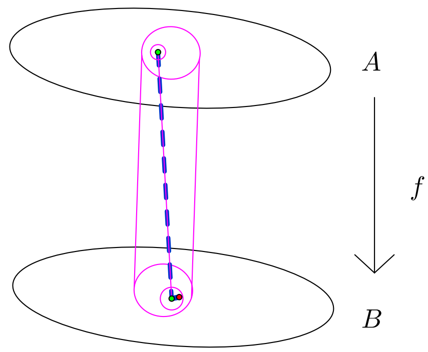
$$A \colon \mathcal{U}, B \colon \mathcal{U}, p \colon \texttt{Id}_{\mathcal{U}}(A,B) \vdash A \simeq B \textbf{ Shape}$$

where the path p in the context does not appear on the RHS. Now, note that the identity map $1_A \equiv \lambda x \colon A.x$ is an equivalence for any $A \colon \mathcal{U}$, which means that we can always find an inhabitant $d(A)$ of $A \simeq A$. But then by `Id-Elim`, we can derive the judgment

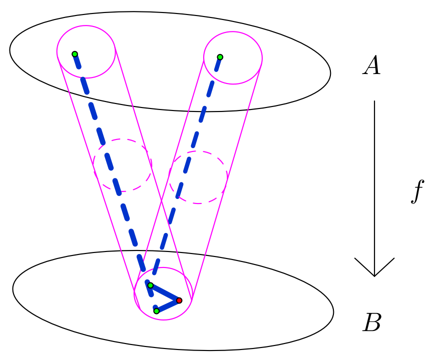
$$A \colon \mathcal{U}, B \colon \mathcal{U}, p \colon \texttt{Id}_{\mathcal{U}}(A,B) \vdash J_d(A,B,p) \colon A \simeq B$$

By Π-`Intro` we then get the desired section

$$\texttt{idtoequiv} =_{\text{df}} \lambda(A,B).\lambda p.J_d(A,B,p) \colon \prod_{A,B \colon \mathcal{U}} \texttt{Id}_{\mathcal{U}}(A,B) \to A \simeq B$$



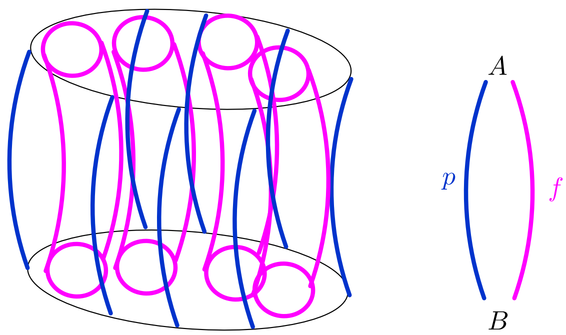
And if this process can be carried out for *every* point of B, viz. if every fiber is contractible, then f would intuitively define exactly a one-to-one correspondence between the points (or better: between the contractible components as defined in (Co)) of A and B. On the other hand, if for example the fiber over a point looked like this



then we would not, intuitively, be able to squash the "double cylinder" down to a line, but only down to a "V" shape. And this, spatially, would capture the fact that f in this case fails to be injective (since it "identifies" two distinct points of A) and therefore does not establish a one-to-one correspondence between points of A and points of B. We conclude, therefore, that the meaning of `isequiv` corresponds to the intuitive meaning of equivalence as stated in (MI). Visually, therefore, an equivalence should be understood as a map that connects two shapes through a series of "cylinders" whose endpoints completely cover them and such that each such "cylinder" can be contracted down to a line.

Now, given a path from A to B (regarded as points of $\mathcal{U}$) we know from (PI) that A and B are indistinguishable. Since the `El`-rule for $\mathcal{U}$ allows us to regard A and B as shapes in their own right, this means that A and B are indistinguishable as shapes. From (MI) it now follows that we can observe an equivalence from A to B. Thus, for any point on the path shape from A to B (regarded as points of $\mathcal{U}$) we can observe a point on the shape of equivalences between A and B (regarded as shapes). By (MS) this gives us a map from the path shape to the shape of equivalences.

Univalence (as we understand it pre-formally) now asserts that this map is itself an equivalence. But by (U) we know that any path can only be turned into an equivalence in a unique way and that conversely every equivalence uniquely determines a path. Therefore, there is a one-to-one correspondence between the points of the path shape and the points of the shape of maps that are equivalences. By (MI) this means that the two shapes are equivalent, as required. Visually, in terms of the above pictures, the right way to think about this correspondence is as follows:



The blue line indicates a path p from A to B in the universe (i.e. a path from A to B when they are regarded as points of the "shape of shapes" $\mathcal{U}$). This path p can be visualized as inducing a "tube" from A to B shown by the "left-leaning" blue curves on the left. An equivalence f between A and B can on the other hand be visualized as inducing all the "right-leaning" cylinders shown in magenta. And univalence then says that each such tube can be uniquely broken down into the cylinders induced by an equivalence and that the cylinders of an equivalence can be uniquely merged into the tube induced by a path. In other words: equivalences are equivalent to paths.

Formal Justification: Univalence can be stated in $\mathcal{U}_\beta$ for $\mathcal{U}_\alpha$ and can be verified to hold in the sense in which the axiomatic presentation would require it to hold. This is recorded as Theorem 3.4.1 in [18].

4 Discussion and Objections

Let me now consider some objections and draw some conclusions. First, one might object that the basic notions of the spatial meaning explanation ("observation", "shape", "symmetry" etc.) are not fundamental enough to deserve to be called "pre-formal". There is not much to say to this other than that they *feel* simple enough to me. As mentioned in the beginning, at some point with genuine meaning explanations, the reader will have to rely on their own cognitive faculties to convince him/herself that the notions being discussed are meaningful. To my mind the notions I have relied on certainly appear to me to be simple enough, or at the very least no less simple that the analogous notions of "stages", "collections" etc. used to justify the axioms of ZFC e.g. by Boolos [7].

Second, even if someone grants that the notions are intuitive, one might still doubt whether the conjunction of the properties (PS)-(U) is plausibly consistent. In words echoing the Frege-Hilbert correspondence: what assurance do we have that these shapes actually exist? There are two things to say in response. Firstly, the formal mathematical model in simplicial sets provides some guarantee that this notion of shape (and "point" and "path" etc.) is at least consistent when it is formalized in set theory. Secondly, shapes as I have understood them have proved themselves to be extremely potent mathematical tools: homotopy theory has been responsible for some major achievements in 20th century mathematics. Their utility in answering mathematical questions beyond their immediate realm of applicability I think also counts in favour of their consistency (or "reality").[18]

Third, one might also object that the basic intuitive notions I have used do not tack on well with the *formal* notions in the set-theoretic models of HoTT. For example, one might protest that the formal notion of a "fibration" does not match my intuitive notion of "observable from an arbitrary point" or that "fiber" does not really correspond to "portion" or that the notion of a "total space" does not match the notion of "juxtaposing shapes". To this I can only reiterate my conviction that these intuitive notions really do tack on to the formally specified ones from algebraic topology. After all, I think it is quite clear that the formal notions originating from algebraic topology, and developed within set theory, are guided by very similar intuitions.

Another big question is the suitability of 1UHoTT as a formal theory for UF amenable to an intuitive justification. As we saw, the justification of $\mathbb{N}$ and $\mathcal{U}$ runs into difficulties. Indeed, there is reason to believe that 1UHoTT will never be fully amenable to the kind of spatial meaning explanation I have sketched in this paper. The reason is that identity types in MLTT remain inescapably "logical". Among other things, this forces us to state univalence as an axiom, instead of proving it as a consequence of properties of other type constructors. The HoTT that is perhaps better-suited to receive my meaning explanation will be something closer to Cubical Type Theory (CTT) where everything other than Π and Σ types are introduced with an explicit spatial intuition, including the path types which are distinct from the identity types of MLTT.[19] As such, in CTT the below-sketched perspective of rules-as-instructions-to-construct-shapes is "written in", rather than tacked on *a posteriori*. In other words, the meaning explanation I have in mind is (implicitly) used in designing the rules of CTT, which would therefore clearly make CTT more amenable to it. In general, to make any progress on this front would require a fully *native* HoTT, i.e. one which in some sense frees itself completely from MLTT (and, more generally, from the family of dependent type theories that were developed with roughly Martin-

[18]A similar argument, discussing the intuitive content of homotopy types, has been put forward by Marquis [22].

[19]Nevertheless, the identity types of standard MLTT (with a judgmental computation rule) are expected to be interpretable into CTT, as described in Section 9.1 of [6].

Löf's meaning explanation in mind). Such a "native HoTT" would be one in which the homotopy interpretation is "written in" from the very beginning, i.e. reflected in the syntax itself. CTT is certainly a step in that direction and so is the type theory outlined in [3].

Finally, there may yet be entirely different approaches to providing meaning explanations to formalizations of the Univalent Foundations. On the one hand, we could imagine maybe a *non-spatial* meaning explanation for HoTT. For example, an alternative approach to the justification of the identity types in HoTT–where types are regarded as "concepts"–is taken by Ladyman and Presnell in [19]. Another idea would be simply to replace the spatial notions I have used in this paper with their "groupoidal" counterparts: for example, I do not think the essential content of (PS)-(U) will change if we replace "shape" by "structure", "path" by "isomorphism" etc.

On the other hand, one can imagine dropping MLTT altogether and working with a simpler "base" formal system, e.g. an expanded version of first-order logic. Such an approach would work analogously to the Boolos-style justifications of ZFC: first, we would justify the rules of a basic system of logic (e.g. $FOL_=$) and then argue for the truth of the non-logical axioms based on a stock of primitive notions (e.g. $\in$). Such an approach would require an expanded $FOL_=$ and then in terms of this expanded system an axiomatization of the universe of ∞-groupoids. I have given my version of what this expanded $FOL_=$ should be in [31]. The system presented there can be thought of as a "first-order logic with *isomorphism*". It is currently work in progress to axiomatize a univalent universe of ∞-groupoids in terms of this system, which would then provide a "first-order" version of UF. And such a system of "first-order UF" could then receive a meaning explanation closer in spirit to the one of Boolos for ZFC.

To conclude: in this paper I have provided a spatial meaning explanation for a relatively simple formalization for the Univalent Foundations which I have called 1UHoTT. If this spatial meaning explanation can be extended to a HoTT that is strong enough to encode all of mathematics as it is currently being practiced, then from this will emerge a picture of the foundations of mathematics in which geometry assumes primacy over other mathematical notions. This signifies a fascinating reversal of the historical trend that led to ZFC, in which mistrust of geometric intuition acted as a primary motivating force. More broadly, this should lead to a reconsideration of the role of geometric thinking not only in mathematics and its foundations, but also perhaps in philosophy and (mathematical) physics.[20] Exploring the limitations and implications of this new picture presents an urgent task.

Appendix: The Rules of 1UHoTT

I will give the full list of rules of the system of 1UHoTT. As in the main text I present them with "**Shape**" instead of "**Type**". My presentation borrows elements from [18] and [34].

[20] This has recently been argued in [11].

Structural Rules

$$\frac{\Gamma \vdash A \ \mathbf{Shape}}{\Gamma \vdash A = A \ \mathbf{Shape}} \ \texttt{Sym-shape-refl} \qquad \frac{\Gamma \vdash A = B \ \mathbf{Shape}}{\Gamma \vdash B = A \ \mathbf{Shape}} \ \texttt{Sym-shape-sym}$$

$$\frac{\Gamma \vdash A = B \ \mathbf{Shape} \qquad \Gamma \vdash B = C \ \mathbf{Shape}}{\Gamma \vdash A = C \ \mathbf{Shape}} \ \texttt{Sym-shape-tran}$$

$$\frac{\Gamma \vdash a \colon A}{\Gamma \vdash a = a \colon A} \ \texttt{Sym-point-refl} \qquad \frac{\Gamma \vdash a = b \colon A}{\Gamma \vdash b = a \colon A} \ \texttt{Sym-point-sym}$$

$$\frac{\Gamma \vdash a = b \colon A \qquad \Gamma \vdash b = c \colon A}{\Gamma \vdash a = c \colon A} \ \texttt{Sym-point-tran}$$

$$\frac{\Gamma \vdash a \colon A \qquad \Gamma \vdash A = B \ \mathbf{Shape}}{\Gamma \vdash a \colon B} \ \texttt{Sym-tran-1} \qquad \frac{\Gamma \vdash a = b \colon A \qquad \Gamma \vdash A = B \ \mathbf{Shape}}{\Gamma \vdash a = b \colon B} \ \texttt{Sym-tran-2}$$

$$\frac{\Gamma \vdash a \colon A \qquad \Gamma, x \colon A, \Delta \vdash b \colon B}{\Gamma, \Delta[a/x] \vdash b[a/x] \colon B[a/x]} \ \texttt{Subst}_1 \qquad \frac{\Gamma \vdash a \colon A \qquad \Gamma, x \colon A, \Delta \vdash b = c \colon B}{\Gamma, \Delta[a/x] \vdash b[a/x] = c[a/x] \colon B[a/x]} \ \texttt{Subst}_2$$

$$\frac{\Gamma \vdash A \ \mathbf{Shape} \qquad \Gamma, \Delta \vdash b \colon B}{\Gamma, x \colon A, \Delta \vdash b \colon B} \ \texttt{Wkg}_1 \qquad \frac{\Gamma \vdash A \ \mathbf{Shape} \qquad \Gamma, \Delta \vdash b = c \colon B}{\Gamma, x \colon A, \Delta \vdash b = c \colon B} \ \texttt{Wkg}_2$$

$$\frac{\begin{array}{c} \vdash A_1 \ \mathbf{Shape} \\ x_1 \colon A_1 \vdash A_2 \ \mathbf{Shape} \\ \vdots \\ x_1 \colon A_1, \ldots, x_{n-1} \colon A_{n-1} \vdash A_n \ \mathbf{Shape} \end{array}}{x_1 \colon A_1, \ldots, x_n \colon A_n \vdash x_j \colon A_j} \ \texttt{Vble}, 1 \leq j \leq n, i \neq j \Rightarrow x_i \neq x_j$$

Rules for type constructors

We omit the rules expressing that logical constructors are preserved under symmetry (i.e. under judgmental equality).

Σ-types

$$\frac{\Gamma \vdash A \ \mathbf{Shape} \qquad \Gamma, x\colon A \vdash B(x) \ \mathbf{Shape}}{\Gamma \vdash \sum_{x\colon A} B(x) \ \mathbf{Shape}} \ \Sigma\texttt{-Form}$$

$$\frac{\Gamma \vdash A \ \mathbf{Shape} \qquad \Gamma, x\colon A \vdash B(x) \ \mathbf{Shape}}{\Gamma, x\colon A, y\colon B(x) \vdash \langle x, y\rangle \colon \sum_{x\colon A} B(x)} \ \Sigma\texttt{-Intro}$$

$$\frac{\begin{array}{c}\Gamma, z\colon \sum_{x\colon A} B(x) \vdash C \ \mathbf{Shape} \\ \Gamma, x\colon A, y\colon B(x) \vdash d(x,y)\colon C(\langle x, y\rangle)\end{array}}{\Gamma, z\colon \sum_{x\colon A} B(x) \vdash \texttt{split}_d(z)\colon C(z)} \ \Sigma\texttt{-Elim}$$

$$\frac{\begin{array}{c}\Gamma, z\colon \sum_{x\colon A} B(x) \vdash C \ \mathbf{Shape} \\ \Gamma, x\colon A, y\colon B(x) \vdash d(x,y)\colon C(\langle x, y\rangle)\end{array}}{\Gamma, x\colon A, y\colon B(x) \vdash \texttt{split}_d(\langle x, y\rangle) = d(x,y)\colon C(\langle x, y\rangle)} \ \Sigma\texttt{-Comp}$$

Π-types

$$\frac{\Gamma \vdash A \ \mathbf{Shape} \qquad \Gamma, x\colon A \vdash B(x) \ \mathbf{Shape}}{\Gamma \vdash \prod_{x\colon A} B(x) \ \mathbf{Shape}} \ \Pi\texttt{-Form}$$

$$\frac{\Gamma, x\colon A \vdash b\colon B(x)}{\Gamma \vdash \lambda(x\colon A).b\colon \prod_{x\colon A} B(x)} \ \Pi\texttt{-Intro} \qquad \frac{\Gamma \vdash f\colon \prod_{x\colon A} B(x) \qquad \Gamma \vdash a\colon A}{\Gamma \vdash f(a)\colon B[a/x]} \ \Pi\texttt{-Elim}$$

$$\frac{\Gamma, x\colon A \vdash b\colon B(x) \qquad \Gamma \vdash a\colon A}{\Gamma \vdash (\lambda(x\colon A).b)(a) = b[a/x]\colon B[a/x]} \ \Pi\texttt{-Comp}$$

$\texttt{Id}$-types

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash a \colon A \qquad \Gamma \vdash b \colon A}{\Gamma \vdash \texttt{Id}_A(a,b) \ \textbf{Shape}} \ \texttt{Id-Form}$$

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash a \colon A}{\Gamma \vdash \texttt{refl}_a \colon \texttt{Id}_A(a,a)} \ \texttt{Id-Intro}$$

$$\frac{\begin{array}{c}\Gamma, x \colon A, y \colon A, p \colon \texttt{Id}_A(x,y) \vdash C \ \textbf{Shape} \\ \Gamma, z \colon A \vdash d(z) \colon C(z,z,\texttt{refl}_z)\end{array}}{\Gamma, x \colon A, y \colon A, p \colon \texttt{Id}_A(x,y) \vdash J_{z.d}(x,y,p) \colon C(x,y,p)} \ \texttt{Id-Elim}$$

$$\frac{\begin{array}{c}\Gamma, x \colon A, y \colon A, p \colon \texttt{Id}_A(x,y) \vdash C \ \textbf{Shape} \\ \Gamma, z \colon A \vdash d(z) \colon C(z,z,\texttt{refl}_z)\end{array}}{\Gamma, x \colon A \vdash J_{z.d}(x,x,\texttt{refl}_x) = d(x) \colon C(x,x,\texttt{refl}_x)} \ \texttt{Id-Comp}$$

$+$-types

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash B \ \textbf{Shape}}{\Gamma \vdash A + B \ \textbf{Shape}} \ +\texttt{-Form}$$

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash B \ \textbf{Shape}}{\Gamma, x \colon A \vdash \texttt{l}(x) \colon A + B} \ +\texttt{-Intro-1} \qquad \frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash B \ \textbf{Shape}}{\Gamma, x \colon A \vdash \texttt{r}(x) \colon A + B} \ +\texttt{-Intro-2}$$

$$\frac{\begin{array}{c}\Gamma, z \colon A + B \vdash C(z) \ \textbf{Shape} \\ \Gamma, x \colon A \vdash d_{\texttt{l}}(x) \colon C(\texttt{l}(x)) \qquad \Gamma, y \colon B \vdash d_{\texttt{r}}(y) \colon C(\texttt{r}(y))\end{array}}{\Gamma, z \colon A + B \vdash \texttt{c}_{d_{\texttt{l}},d_{\texttt{r}}}(z) \colon C(z)} \ +\texttt{-Elim}$$

$$\frac{\begin{array}{c}\Gamma, z \colon A + B \vdash C(z) \ \textbf{Shape} \\ \Gamma, x \colon A \vdash d_{\texttt{l}}(x) \colon C(\texttt{l}(x)) \qquad \Gamma, y \colon B \vdash d_{\texttt{r}}(y) \colon C(\texttt{r}(y))\end{array}}{\Gamma, x \colon A \vdash \texttt{c}_{d_{\texttt{l}},d_{\texttt{r}}}(\texttt{l}(x)) = d_{\texttt{l}}(x) \colon C(\texttt{l}(x))} \ +\texttt{-Comp-1}$$

$$\frac{\begin{array}{c}\Gamma, z \colon A + B \vdash C(z) \ \textbf{Shape} \\ \Gamma, x \colon A \vdash d_{\texttt{l}}(x) \colon C(\texttt{l}(x)) \qquad \Gamma, y \colon B \vdash d_{\texttt{r}}(y) \colon C(\texttt{r}(y))\end{array}}{\Gamma, y \colon B \vdash \texttt{c}_{d_{\texttt{l}},d_{\texttt{r}}}(\texttt{r}(y)) = d_{\texttt{r}}(y) \colon C(\texttt{r}(y))} \ +\texttt{-Comp-2}$$

0-type

$$\frac{}{\Gamma \vdash \mathbf{0}\ \textbf{Shape}}\ \texttt{0-Form} \qquad \frac{\Gamma, x\colon \mathbf{0} \vdash C(x)\ \textbf{Shape}}{\Gamma, x\colon \mathbf{0} \vdash \texttt{efq}\colon C(x)}\ \texttt{0-Elim}$$

1-type

$$\frac{}{\Gamma \vdash \mathbf{1}\ \textbf{Shape}}\ \texttt{1-Form} \qquad \frac{}{\Gamma \vdash *\colon \mathbf{1}}\ \texttt{1-Intro}$$

$$\frac{\Gamma x\colon \mathbf{1} \vdash C(x)\ \textbf{Shape} \qquad \Gamma \vdash d\colon C(*)}{\Gamma, x\colon \mathbf{1} \vdash \texttt{rec}_d(x)\colon C(x)}\ \texttt{1-Elim}$$

$$\frac{\Gamma x\colon \mathbf{1} \vdash C(x)\ \textbf{Shape} \qquad \Gamma \vdash d\colon C(*)}{\Gamma \vdash \texttt{rec}_d(*) = d\colon C(*)}\ \texttt{1-Comp}$$

$\mathbb{N}$-type

$$\frac{}{\Gamma \vdash \mathbb{N}\ \textbf{Shape}}\ \mathbb{N}\texttt{-Form} \qquad \frac{}{\Gamma, \vdash 0\colon \mathbb{N}}\ \mathbb{N}\texttt{-Intro-1} \qquad \frac{\Gamma \vdash n\colon \mathbb{N}}{\Gamma \vdash s(n)\colon \mathbb{N}}\ \mathbb{N}\texttt{-Intro-2}$$

$$\frac{\begin{array}{c}\Gamma, x\colon \mathbb{N} \vdash C\ \textbf{Shape} \\ \Gamma \vdash c_0\colon C[0/x] \qquad \Gamma, x\colon \mathbb{N}, y\colon C \vdash c_s\colon C[s(x)/x] \qquad \Gamma \vdash n\colon \mathbb{N}\end{array}}{\Gamma \vdash \texttt{ind}_{c_0, c_s, n}\colon C[n/x]}\ \mathbb{N}\texttt{-Elim}$$

$$\frac{\Gamma, x\colon \mathbb{N} \vdash C\ \textbf{Shape} \qquad \Gamma \vdash c_0\colon C[0/x] \qquad \Gamma, x\colon \mathbb{N}, y\colon C \vdash c_s\colon C[s(x)/x]}{\Gamma \vdash \texttt{ind}_{c_0, c_s, 0} = c_0\colon C[0/x]}\ \mathbb{N}\texttt{-Comp-1}$$

$$\frac{\begin{array}{c}\Gamma, x\colon \mathbb{N} \vdash C\ \textbf{Shape} \\ \Gamma \vdash c_0\colon C[0/x] \qquad \Gamma, x\colon \mathbb{N}, y\colon C \vdash c_s\colon C[s(x)/x] \qquad \Gamma \vdash n\colon \mathbb{N}\end{array}}{\Gamma \vdash \texttt{ind}_{c_0, c_s, s(n)} = c_s\colon C[s(n)/x]}\ \mathbb{N}\texttt{-Comp-2}$$

Rules for the Universe

$$\frac{}{\vdash \mathcal{U}\ \textbf{Shape}}\ \texttt{Uni-Form} \qquad \frac{}{x\colon \mathcal{U} \vdash \texttt{El}(x)\ \textbf{Shape}}\ \texttt{Uni-El}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma \vdash b\colon \mathcal{U}}{\Gamma \vdash u_{a,b}\colon \texttt{isequiv}(\texttt{idtoequiv}_{a,b})}\ \texttt{Univalence}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma, x\colon \texttt{El}(a) \vdash b(x)\colon \mathcal{U}}{\Gamma \vdash \pi(a, x.b(x))\colon \mathcal{U}}\ \texttt{Uni-}\Pi\texttt{-1}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma, x\colon \texttt{El}(a) \vdash b(x)\colon \mathcal{U}}{\Gamma \vdash \texttt{El}(\pi(a, x.b(x))) = \prod_{x\colon \texttt{El}(a)} \texttt{El}(b(x))\ \textbf{Shape}}\ \texttt{Uni-}\Pi\texttt{-2}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma, x\colon \texttt{El}(a) \vdash b(x)\colon \mathcal{U}}{\Gamma \vdash \sigma(a, x.b(x))\colon \mathcal{U}}\ \texttt{Uni-}\Sigma\texttt{-1}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma, x\colon \texttt{El}(a) \vdash b(x)\colon \mathcal{U}}{\Gamma \vdash \texttt{El}(\sigma(a, x.b(x))) = \sum_{x\colon \texttt{El}(a)} \texttt{El}(b(x))\ \textbf{Shape}}\ \texttt{Uni-}\Sigma\texttt{-2}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma \vdash b\colon \texttt{El}(a) \qquad \Gamma \vdash c\colon \texttt{El}(a)}{\Gamma \vdash \texttt{id}_a(b, c)\colon \mathcal{U}}\ \texttt{Uni-Id-1}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma \vdash b\colon \texttt{El}(a) \qquad \Gamma \vdash c\colon \texttt{El}(a)}{\Gamma \vdash \texttt{El}(\texttt{id}_a(b, c)) = \texttt{Id}_{\texttt{El}(a)}(b, c)\ \textbf{Shape}}\ \texttt{Uni-Id-2}$$

$$\frac{}{\vdash \texttt{z}\colon \mathcal{U}}\ \texttt{Uni-}\mathbf{0}\texttt{-1} \qquad \frac{}{\vdash \texttt{El}(\texttt{z}) = \mathbf{0}\ \textbf{Shape}}\ \texttt{Uni-}\mathbf{0}\texttt{-2}$$

$$\frac{}{\vdash \texttt{s}\colon \mathcal{U}}\ \texttt{Uni-}\mathbf{1}\texttt{-1} \qquad \frac{}{\vdash \texttt{El}(\texttt{s}) = \mathbf{1}\ \textbf{Shape}}\ \texttt{Uni-}\mathbf{1}\texttt{-2}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma \vdash b\colon \mathcal{U}}{\Gamma \vdash a \oplus b\colon \mathcal{U}}\ \texttt{Uni-}+\texttt{-1}$$

$$\frac{\Gamma \vdash a\colon \mathcal{U} \qquad \Gamma \vdash b\colon \mathcal{U}}{\Gamma \vdash \texttt{El}(a \oplus b) = \texttt{El}(a) + \texttt{El}(b)\ \textbf{Shape}}\ \texttt{Uni-}+\texttt{-2}$$

$$\frac{}{\vdash \texttt{N}\colon \mathcal{U}}\ \texttt{Uni-}\mathbb{N}\texttt{-1} \qquad \frac{}{\vdash \texttt{El}(\texttt{N}) = \mathbb{N}\ \textbf{Shape}}\ \texttt{Uni-}\mathbb{N}\texttt{-2}$$

References

[1] Aczel, P.: The Type Theoretic Interpretation of Constructive Set Theory. In: L.P. A. MacIntrye, J. Paris (eds.) Logic Colloquium ’77 (Proc. Conf., Wroclaw, 1977), Stud. Logic Foundations Math., vol. 96, pp. 55–66. North-Holland, Amsterdam-New York (1978)

[2] Altenkirch, T., McBride, C.: Towards observational type theory. Manuscript, available online (2006)

[3] Angiuli, C., Harper, R., Wilson, T.: Computational Higher Type Theory I: Abstract Cubical Realizability. arXiv preprint arXiv:1604.08873 (2016)

[4] Awodey, S.: Structuralism, Invariance and Univalence. Philosophia Mathematica **22 (1)**, 1–11 (2014)

[5] Awodey, S., Warren, M.A.: Homotopy theoretic models of identity types. Math. Proc. Camb. Phil. Soc. **146**(45), 45–55 (2009)

[6] Bezem, M., Coquand, T., Huber, S.: A model of type theory in cubical sets. In: 19th International Conference on Types for Proofs and Programs (TYPES 2013), vol. 26, pp. 107–128 (2014)

[7] Boolos, G.: The iterative conception of set. The Journal of Philosophy **68**(8), 215–231 (1971)

[8] Burgess, J.: Rigor and Structure. Oxford University Press, Oxford (2014)

[9] Cartmell, J.: Generalized algebraic theories and contextual categories. Annals of Pure and Applied Logic **32**, 209–243 (1986)

[10] Cohen, C., Coquand, T., Huber, S., Mörtberg, A.: Cubical type theory: a constructive interpretation of the univalence axiom (2015). `https://www.math.ias.edu/~amortberg/papers/cubicaltt.pdf`

[11] Corfield, D.: Reviving the philosophy of geometry (2015). URL `http://philsci-archive.pitt.edu/11809/`

[12] Corfield, D.: Expressing “the structure of” in homotopy type theory (2016). URL `http://philsci-archive.pitt.edu/11862/`

[13] Dybjer, P.: Internal type theory. In: Types for Proofs and Programs, pp. 120–134. Springer (1995)

[14] Dybjer, P.: Program testing and the meaning explanations of intuitionistic type theory (2012). Available at `http://www.cse.chalmers.se/~peterd/papers/MartinLofFestschrift.pdf`

[15] Dybjer, P., Lindström, S., Palmgren, E., Sundholm, B.G.: Epistemology versus ontology: Essays on the philosophy and foundations of mathematics in honour of per Martin-Löf, vol. 27. Springer Science & Business Media (2012)

[16] Griffor, E., Rathjen, M.: The strength of some martin-löf type theories. Archive for Mathematical Logic **33**(5), 347–385 (1994)

[17] Hellman, G., Shapiro, S.: The Classical Continuum Without Points. Review of Symbolic Logic **6**(3), 488–512 (2013)

[18] Kapulkin, K., Lumsdaine, P., Voevodsky, V.: The simplicial model of univalent foundations. arXiv:1211.2851v2 (2014)

[19] Ladyman, J., Presnell, S.: Identity in Homotopy Type Theory, Part I: The Justification of Path Induction. Philosophia Mathematica (2015)

[20] Lawvere, W.: An elementary theory of the category of sets. Reprints in Theory and Applications of Categories **12**, 1–35 (2005)

[21] Linnebo, Ø., Pettigrew, R.: Category theory as an autonomous foundation. Philosophia Mathematica pp. 1–30 (2011)

[22] Marquis, J.P.: Mathematical forms and forms of mathematics: leaving the shores of extensional mathematics. Synthese **190**(12), 2141–2164 (2013)

[23] Martin-Löf, P.: Constructive mathematics and computer programming. Studies in Logic and the Foundations of Mathematics **104**, 153–175 (1982)

[24] Martin-Löf, P.: Intuitionistic type theory. Bibliopolis (1984)

[25] Martin-Löf, P.: Truth of a Proposition, Evidence of a Judgement, Validity of a Proof. Synthese (73), 407–420 (1987)

[26] Martin-Löf, P.: On the meanings of the logical constants and the justifications of the logical laws. Nordic journal of philosophical logic **1**(1), 11–60 (1996)

[27] Pelayo, Á., Warren, M.: Homotopy type theory and Voevodsky's univalent foundations. Bulletin of the American Mathematical Society **51**(4), 597–648 (2014)

[28] Rodin, A.: On constructive axiomatic method (2014). Available at `http://philsci-archive.pitt.edu/10986/`

[29] Setzer, A.: Well-ordering proofs for martin-löf type theory. Annals of Pure and applied Logic **92**(2), 113–159 (1998)

[30] Shulman, M.: Homotopy type theory: A synthetic approach to higher equalities. In: E. Landry (ed.) Categories for the Working Philosopher (2016)

[31] Tsementzis, D.: Homotopy Model Theory I: Syntax and Semantics. arXiv preprint arXiv:1603.03092 (2016)

[32] Tsementzis, D.: Univalent foundations as structuralist foundations. Synthese pp. 1–35 (2016). DOI 10.1007/s11229-016-1109-x. URL http://dx.doi.org/10.1007/s11229-016-1109-x

[33] Tsementzis, D.: What is a higher-level set? Philosophia Mathematica (2017). https://doi.org/10.1093/philmat/nkw032

[34] Univalent Foundations Program: Homotopy type theory: Univalent foundations of mathematics (2013). http://homotopytypetheory.org/book

[35] Voevodsky, V.: Foundations of mathematics and homotopy theory (2006). https://www.math.ias.edu/vladimir/sites/math.ias.edu.vladimir/files/VV%20Slides.pdf

[36] Voevodsky, V.: The equivalence axiom and univalent models of type theory. Talk at CMU (2010)

[37] Voevodsky, V.: Univalent foundations project (2010). http://www.math.ias.edu/vladimir/files/univalent_foundations_project.pdf

[38] Voevodsky, V.: Subsystems and regular quotients of C-systems. arXiv 1406.5389, submitted p. 1–11 (2014). URL http://arxiv.org/abs/1406.7413

[39] Voevodsky, V.: Univalent Foundations (lecture at the IAS) (2014). https://www.math.ias.edu/vladimir/sites/math.ias.edu.vladimir/files/2014_IAS.pdf

[40] Voevodsky, V., Ahrens, B., Grayson, D., et al.: *UniMath*: Univalent Mathematics. Available at https://github.com/UniMath

[41] Warren, M.A.: Homotopy theoretic aspects of constructive type theory. Ph.D. thesis, Carnegie Mellon University (2008)