A Meaning Explanation for HoTT*

Dimitris Tsementzis

March 12, 2018

Abstract

In the Univalent Foundations of mathematics spatial notions like "point" and "path" are primitive, rather than derived, and all of mathematics is encoded in terms of them. A Homotopy Type Theory is any formal system which realizes this idea. In this paper I will focus on the question of whether a Homotopy Type Theory (as a formalism for the Univalent Foundations) can be justified intuitively as a theory of shapes in the same way that ZFC (as a formalism for set-theoretic foundations) can be justified intuitively as a theory of collections. I first clarify what such an "intuitive justification" should be by distinguishing between formal and pre-formal "meaning explanations" in the vein of Martin-Löf. I then go on to develop a pre-formal meaning explanation for HoTT in terms of primitive spatial notions like "shape", "path" etc.

Introduction

The Univalent Foundations (UF) of mathematics take the point of view that the basic objects of mathematics are *shapes* rather than *sets*. UF is usually formalized in terms of formal theories called Homotopy Type Theories (HoTT(s)).[1]

Set theory, usually formalized as ZFC, is the currently accepted foundation of mathematics by the community of mathematicians.[2] A pressing issue is therefore to articulate the relationship between set-theoretic foundations and

*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 Harry Crane for supporting, encouraging and carefully dissecting the ideas in this paper, Harvey Friedman for his penetrating criticism of the "new foundations" that led to a significant sharpening of the particular aims of this paper, Bob Harper for his many comments, questions and for patiently clarifying to me his views on meaning explanations, and Favonia (Kuen-Bang Hou) for teaching me almost everything I know about the meaning explanation underlying CHTT. 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 [15, 27, 29, 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, 31–33]. For some philosophical issues associated to HoTT see [11, 12, 19, 28].

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

the Univalent Foundations, as well as more specifically that of ZFC and HoTT as their respective formalizations.[3]

In this paper we will focus on the following question: can we understand the basic objects of UF ("shapes") in a way that is *autonomous* from set theory? One way to make the question of autonomy precise is as follows: can the rules of some HoTT be justified *pre-formally* in terms of notions that are independent from naive set theory?[4]

To answer this question requires describing some such pre-formal notions, and using them to supply the syntax of HoTT with meaning in terms of which its rules are justified (or, as I will say, *visualized*). 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."

1 On Meaning Explanations

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

$$\frac{s_1 \qquad s_2 \qquad \dots \qquad s_n}{t}\ R$$

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".

[3]Some facts are already known about the technical relationship between ZFC and HoTT. For example, the simplicial model of univalence [18] gives an interpretation of HoTT into ZFC supplemented with two inaccessible cardinals and Cubical Type Theory (CTT) [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]) known to be equiconsistent with the system BZC.

[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].

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 $\wedge$-introduction 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.[5] 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. 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

[5]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.

formal system in some other formal system. 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 want to do in this paper is to provide 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 resolutely not the kind of formal interpretation of HoTT in set theory sketched in [6, 18]. My efforts here more closely compare to Computational Higher Type Theory (CHTT) [3, 17] which seeks (among other things) to provide computational meaning explanations to HoTT-style systems in terms of realizability semantics - although the CHTT developers also concede that their “meaning explanation” is a *mathematical* one. 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 A Meaning Explanation for HoTT

HoTT as currently understood is based on Martin-Löf Type Theory (MLTT) [24].[6] Martin-Löf gave a thorough meaning explanation for MLTT which has guided the development of MLTT. This meaning explanation understood the basic syntactic elements of MLTT (“judgments”) in terms of notions of computation and evaluation to canonical values, mixed in with a BHK-style interpretation of the logical connectives and quantifiers.[7] For example, the meaning of the judgment

$$A\ \mathbf{Type}$$

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

$$a : A$$

was that a is a program that satisfies the specification of A or, equivalently, which can carry out the task specified by A.[8]

[6] This does not mean that any formalization for UF will be based on MLTT. UF is no more tied to MLTT than set-theoretic foundations are tied to Russell’s Type Theory. I take up this point in the conclusion.

[7] For more details see [23, 26] as well as remarks in the original [24]. A good exposition can be found in [13].

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

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 any 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.[9] 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, we have to provide 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. [10]

To do so, we will now describe a realm of shapes and use this pre-formal description of shapes to give meaning to the syntax and to justify the rules of a specific HoTT. The syntax and rules for our HoTT have been given in full in the appendix. It consists of:

1. The usual syntax of MLTT and associated structural rules.

2. Rules for the following constructors: Total ("Σ"), Map ("Π"), Path ("Id"), $\mathbf{0}$, $\mathbf{1}$ and $\mathbb{N}$.

3. A sequence of cumulative universes $\mathcal{U}_i$ closed under all the above constructors for $i \in \omega$.[11]

4. Propositional resizing and univalence for each universe.

From now on I will use the term "HoTT" to refer to this specific formal system.[12]

[9] 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 [13], p. 8 for further discussion on this point and [17], p. 18 for more on the incompatibility of UIP with univalence.

[10] 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 a "shape 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.

[11] The notation ω here is used to emphasize that the universes are ordered, and there could be countably infinite of them, as well as in order to distinguish the indexing of the universes from the notation $\mathbb{N}$. The use of set-theoretic language here is purely conventional, occurs at the meta-theory level, and carries no commitments.

[12] UniMath [40] is the closest formal realization of my HoTT, 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, HoTT 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). 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) or Computational Higher Type Theory (CHTT) might prove better-suited to this task.

2.1 The Meaning of the Syntax

We start with a space of shapes. Everything in space is a shape. All shapes exist in space. Shapes have points.

We access shapes and their points through *observations*. There are four fundamental observations we can make. We can observe a shape

$$A \textbf{ Shape}$$

We can observe a point on a shape

$$a : A$$

To observe a shape is to observe a certain presentation of a shape (to observe it from a certain angle, as it were). Thus, the same shape can be presented to us in different ways and two such presentations are symmetric if they are presentations of the same shape

$$A \equiv B \textbf{ Shape}$$

Similarly, the same point can be presented to us in different ways (from different viewpoints, as it were)

$$a \equiv b : A$$

We can write $\mathcal{O}$ for any of the four types of observation.

Observations can be made from *viewpoints*. A viewpoint consists of the points of a shape and given any viewpoint Γ and shape A we write

$$\Gamma, x : A \textbf{ view}$$

for the viewpoint which extends Γ by the points of A.

If Γ is a viewpoint then we write

$$\Gamma \vdash \mathcal{O}$$

to express that the observation $\mathcal{O}$ is made from viewpoint Γ. Thus, the meaning of an expression $\Gamma \vdash \mathcal{O}$ for some observation $\mathcal{O}$ is to be understood as

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

For example, the expression

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

means

"From any point x of A, we can observe a certain shape B"

This concludes the explanation of the meaning of the syntax of our HoTT. To summarize, what in MLTT are called "judgments" we will understand as *observations* and what in MLTT are called "contexts" we will understand as *viewpoints*.[13] There are four types of observations, corresponding to the four types of judgment in MLTT:

Notation	Meaning
$a : A$	a is a **point** of A
$a \equiv a' : A$	a and a' are **symmetric** as points of A
A **Shape**	A is a **shape**
$A \equiv A'$ **Shape**	A and A' are **symmetric** as shapes

Each such observation can be "relativized" to a viewpoint. To observe a shape from a viewpoint is then to observe only a "portion" of the "whole" shape, namely the portion that is visible from that specific viewpoint. What we observe of a shape, therefore, can *depend* on a viewpoint.[14]

Overall, we have four primitive notions that give meaning to the syntax of MLTT, on which the whole proposed meaning explanation for HoTT rests: **viewpoint** (replacing "context"), **point** (replacing "term") **shape** (replacing "type") and **symmetric** (replacing "judgmentally equal")

2.2 The Visualization of the Rules

Given the interpretation of judgments as observations and contexts as viewpoints, we will now justify the rules of HoTT. Before we begin, I want to introduce a neologism, which explains the title of this subsection: instead of saying that we will *justify* the rules of HoTT, I will instead say we will *visualize* these rules.

The reason for using "visualize" instead of "justify" is that the rules of the specific HoTT that I will visualize/justify are not to be understood merely as rules asserting the mere *preservation* of some property from what is above the line to what is below the line (e.g. truth-preservation). Rather, they are to be understood as describing certain *operations* (or *constructions*) in space. The rules of HoTT are not assertions about the preservation of *properties*; they are descriptions of the existence of *structure*. Thus, a rule of HoTT is not to be *justified* merely as a statement about what observations *can* be made (given certain observation already made), but rather it will be *visualized* by *how* these observations are made (given how certain observations have already been made). In slogan form: *properties can be justified, but structure must be visualized.*[15]

In what we say below, therefore, one should not ask "Is the rule justified?" but rather "Can I visualize it?" I believe this distinction is fundamental, and

[13] 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.

[14] This is how "type dependency" is parsed in our meaning explanation.

[15] This slogan reflects the formal fact that in the formal models of the kind of type theories on which HoTT is based (e.g. models in terms of contextual categories) the rule for type constructors describe *structure* on these formal models (e.g. Σ-type structure on a contextual category), rather than properties. My decision to adopt the neologism here described is just a pre-formal reflection of this fact.

justifies the potentially disconcerting (and annoying) usage of sentences like "Fact P *visualizes* rule R" as replacements (indeed, generalizations) for the sentence "Fact P *justifies* rule R".

2.2.1 Structural Rules

There is a basic or "neutral" viewpoint, from which observations can begin, which is how we visualize the following rule:

$$\frac{}{\bullet\ \textbf{view}}\ \bullet\text{ - }\textbf{view}$$

Suppose we are standing on a viewpoint Γ and from it we can observe a shape. As it were, we expand any viewpoint to include any shape we can observe from it, and this is how we visualize the following rule:

$$\frac{\Gamma\ \textbf{view} \qquad \Gamma \vdash A\ \textbf{Shape}}{\Gamma, x : A\ \textbf{view}}\ \mathsf{ext}\text{ - }\textbf{view}$$

Intuitively it tells us where exactly we can "move around" on the new shape we have added to our viewpoint in order to make new observations.

Now, given any viewpoint that contains a shape A we can always observe a point on A. Clearly, if we are "standing" on a viewpoint we can observe any of its points. And this is how we visualize the following rule:

$$\frac{\Gamma, x : A, \Delta\ \ \textbf{view}}{\Gamma, x : A, \Delta \vdash x : A}\ \mathsf{ax}\text{ - }\textbf{view}$$

The structural rules for symmetry observations (`Sym-shape-sym`, `Sym-shape-tran` etc.) are visualized 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

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

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 we have observed them from a different perspective. 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.

2.2.2 Rules for Shape Constructors

The rules for each shape constructor must now be visualized as describing how to construct new shapes obtained from previously-obtained shapes. For example, a formation rule answers the question "When can we observe X?" and an introduction rule answers the question "When can we observe a point of X?" In what follows I will generally forgo mention of the (always assumed) implicit viewpoint Γ. The reader may assume that all the visualizations below take place relative to a given viewpoint and this does not alter their essential content.

Map-shapes as shapes of maps

If a shape can be observed from another shape there is a shape of maps of the latter shape into the former. A map is an assignment of the points of a shape to the points of the shape it maps to. So if we can observe a shape A and from any x on A we can observe a shape $B(x)$ then we introduce the symbol

$$\underset{x:\,A}{\mathsf{Map}}\ B(x)$$

to represent the shape of maps from A to the shapes $B(x)$ visible from each point x of A. This is the *map shape* or *shape of sections*, which allows us to visualize the following rule

$$\frac{\Gamma \vdash A\ \textbf{Shape} \qquad \Gamma, x\colon A \vdash B(x)\ \textbf{Shape}}{\Gamma \vdash \underset{x:\,A}{\mathsf{Map}}\ B(x)\ \textbf{Shape}}\ \texttt{Map-Form}$$

Whenever A and B can be observed from the same viewpoint (i.e. whenever there is no "dependency") we will also simply write

$$A \to B$$

for the map shape from A to B.

Now assume that for any point x on A we can observe a point $b(x)$ on the portion $B(x)$. But 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 $B(x)$. Such a map is exactly a point of $\underset{x:\,A}{\mathsf{Map}}\ B(x)$ as just described, which allows us to visualize the following rule

$$\frac{\Gamma, x\colon A \vdash b(x)\colon B(x)}{\Gamma \vdash \lambda(x\colon A).b(x)\colon \underset{x:\,A}{\mathsf{Map}}\ B(x)}\ \texttt{Map-Intro}$$

in which the notation $\lambda(x : A).b(x)$ is understood as the map which takes a point x on A to the point $b(x)$ on $B(x)$ we observe from it.

Now assume we can observe a point f on the map shape $\underset{x:\,A}{\mathsf{Map}}\ B(x)$, i.e. a map from A to portions of B and also a point a on A. But since f is a map we know there we can observe the point $f(a)$ on the shape $B(a)$ visible from a, which visualizes the following rule:

$$\frac{\Gamma \vdash f \colon \underset{x\colon A}{\mathsf{Map}}\, B(x) \qquad \Gamma \vdash a \colon A}{\Gamma \vdash f(a) \colon B[a/x]}\ \texttt{Map-Elim}$$

Now 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 (i.e. a map) on $\underset{x\colon A}{\mathsf{Map}}\, B(x)$ and then this point will give us a corresponding point on $B(a)$. On the observation $\Gamma \vdash a : A$ we know that from a some point will be observable on $B(a)$. These two points are clearly the same point, which visualizes the following rule:

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

This completes the visualization for the rules of map shapes.

Total-shapes as total shapes

If a shape can be observed from the viewpoint of another shape, there is a shape whose points are pairs consisting of a point of the first shape and a point on the shape observable from that point. This is the *total shape* which visualizes the following rule:

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

The points of the total shape are pairs of points, where the second pair is a point on the shape observable from the first. So if we can observe a point a on a shape A and a point b on the shape $B[a/x]$ observable from a then we can observe a point on the total shape. This visualizes the following rule:

$$\frac{\begin{array}{c}\Gamma \vdash A\ \textbf{Shape} \qquad \Gamma, x \colon A \vdash B(x)\ \textbf{Shape} \\ \Gamma \vdash a \colon A \qquad \Gamma \vdash b \colon B[a/x]\end{array}}{\Gamma \vdash \langle a, b \rangle \colon \underset{x\colon A}{\mathsf{Total}}\, B(x)}\ \texttt{Total-Intro}$$

Every point of the total shape must be visualized exactly as a pair of points. Therefore, for each such point on the total shape that we have observed, we can *project* it onto the points on the shapes that make up the total shape. In other words, we can always observe the “shadow” of a pair on the shapes from which that pair was formed. This is how we visualize the following rules:

$$\frac{\Gamma \vdash z \colon \underset{x\colon A}{\mathsf{Total}}\, B(x)}{\Gamma \vdash \mathsf{pr1}(z) : A}\ \mathsf{Total\text{-}pr1} \qquad \frac{\Gamma \vdash z \colon \underset{x\colon A}{\mathsf{Total}}\, B(x)}{\Gamma \vdash \mathsf{pr2}(z) : B[\mathsf{pr1}(z)/x]}\ \mathsf{Total\text{-}pr2}$$

These "projected" (or "shadow") points are visualized as exactly the points the pair is made of and conversely any point of the total shape is symmetric to the pair of points that one obtains by "projecting". This is how we visualize the following rules:

$$\frac{\Gamma \vdash a\colon A \qquad \Gamma \vdash b\colon B[a/x]}{\Gamma \vdash \mathsf{pr1}(\langle a, b\rangle) \equiv a : A \qquad \Gamma \vdash \mathsf{pr2}(\langle a, b\rangle) \equiv b : B[a/x]}\ \texttt{Total-comp-1}$$

$$\frac{\Gamma \vdash z\colon \underset{x\colon A}{\mathsf{Total}}\ B(x)}{\Gamma \vdash z \equiv \langle \mathsf{pr1}(z), \mathsf{pr2}(z)\rangle : \underset{x\colon A}{\mathsf{Total}}\ B(x)}\ \texttt{Total-comp-2}$$

This completes the visualization for the rules of total shapes.

1-shapes as singleton shapes

A point is a shape, which means that we can visualize a shape with a single point. This is the *singleton shape*, which visualizes the following rules:

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

Assume we have observed a point a on a shape A. Then we can observe a map from $\mathbf{1}$ to A that takes the unique point of $\mathbf{1}$ to a (since a map is an assignment of points to points and there is only one point in $\mathbf{1}$ that we can assign to A). This visualizes the following two rules:

$$\frac{\Gamma \vdash A\ \textbf{Shape} \qquad \Gamma \vdash a\colon A}{\Gamma \vdash !_a\colon \mathbf{1} \to A}\ \texttt{1-Elim} \qquad \frac{\Gamma \vdash A\ \textbf{Shape} \qquad \Gamma \vdash a\colon A}{\Gamma \vdash !_a(a) \equiv a\colon A}\ \texttt{1-Comp}$$

Path-shapes as shapes of paths between points

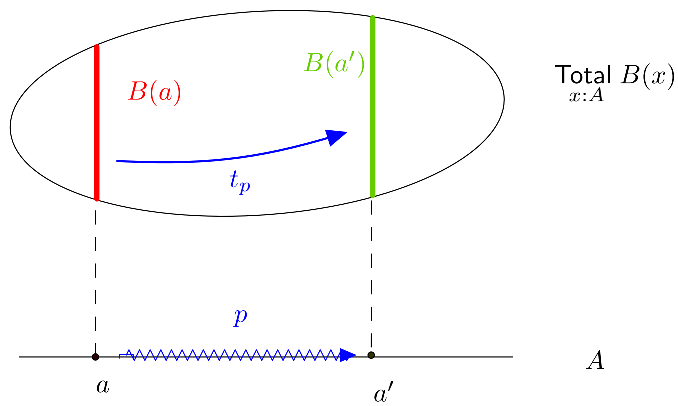
Points on a shape can be connected by a path. The paths between two points also form a shape (since everything is a shape) and the points of that shape are the paths. This is the *path shape*, which visualizes the following rule:

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

We can always observe the *trivial* or *reflexivity* path from a point to itself, which is the path that remains constant, i.e. the "non-path" from a point to itself, and which visualizes the following rule:

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

Now, suppose that we can observe two points a, a' on a shape A and a path $p : \mathsf{Path}_A(a, a')$ between them, and suppose that from the viewpoint of A we observe shapes B. Then the path p "lifts" to a map t_p between the shape $B(a)$ we observe from a and the shape $B(a')$ we observe from b. We can visualize this as follows:



The spatial intuition here is that "above" we map via t_p the red shape $B(a)$ on the left to the green shape $B(a')$ on the right by simply tracing the points of the former onto the points of the latter, along p "below". This is exactly how we visualize the following rule:

$$\frac{\begin{array}{c}\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, x : A \vdash B \ \textbf{Shape} \\ \Gamma \vdash a : A \qquad \Gamma \vdash a' : A \\ \Gamma \vdash p : \mathsf{Path}_A(a, a')\end{array}}{\Gamma \vdash t_p \colon B[a/x] \to B[a'/x]} \ \texttt{Path-transport}$$

We call the map t_p the *transport (map) along* p. Clearly the transport along the trivial path leaves any point exactly where it was, since we trace each observed point to itself. This is how we visualize the following rule:

$$\frac{\begin{array}{c}\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, x : A \vdash B \ \textbf{Shape} \\ \Gamma \vdash a : A \qquad \Gamma \vdash b : B[a/x]\end{array}}{\Gamma \vdash t_{\mathsf{refl}_a}(b) \equiv b : B[a/x]} \ \texttt{Path-transport-sym}$$

Visualizing the paths between points on shapes is fundamental. How these paths behave is exactly what determines the notion of a shape that we are visualizing. The key spatial intuition here is that we can make no observation from a point on a shape that cannot also be "carried" along a path. Thus, any point on a shape can be "deformed" to any other point to which it is connected by a path. Every observation we make on a shape is "path-invariant" – and as such, we must visualize shapes as

> [...] some sort of vague space, which can be very severely deformed, mapped one to another, and all the while the specific space is not important, but only the space up to deformation. If we really want

> to return to discrete objects, we see continuous components, the pieces whose form or even dimension does not matter. Earlier, all these spaces were thought of as Cantor sets with topology, their maps were Cantor maps, some of them were homotopies that should have been factored out, and so on. [...] [P]oints become continuous components, or attractors [...] almost from the start. (Y. Manin in [14], p.1274)

Keeping this visualization in mind, we can now make meaningful definitions of a contractible shape and of the notion of an equivalence of shapes.

Contractible Shapes

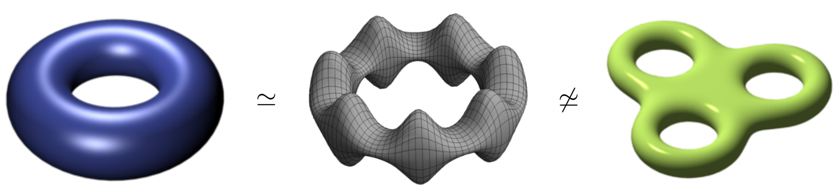
A shape is *contractible* if it can be continuously deformed down to a point. In other words, a shape is contractible if up to deformation it is indistinguishable from a point. More precisely, we can define a shape to be contractible if there is a point on it such that every other point can be *continuously* connected to it by a path. In other words: a shape is contractible if it has only one point *up to a path*. Thus, in order to observe that a shape is contractible we need to observe that there is such a point (a "center of contraction") and a map that takes any other point to a path connecting them. To observe such a pair is exactly to observe a point on a certain total shape, which is how we visualize the following definition

$$\mathsf{iscontractible}(A) =_{\mathrm{df}} \underset{a\colon A}{\mathsf{Total}}\ \underset{x\colon A}{\mathsf{Map}}\ \mathsf{Path}_A(x, a)$$

Thus, to observe that a shape A is contractible is formalized in HoTT as the observation of a point on $\mathsf{iscontractible}(A)$.

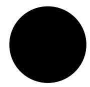
Equivalence of Shapes

Two shapes are equivalent if they can be continuously deformed into one another by a map without tearing or gluing.[16] Thus, a deformed torus is equivalent to a smooth torus and not equivalent to a torus with many holes

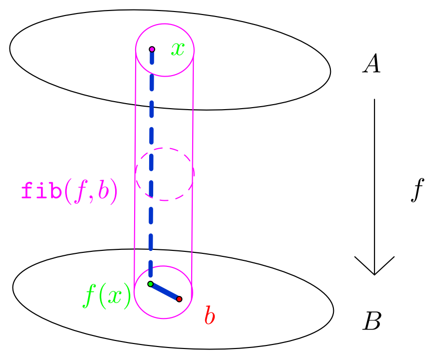


and a disc is equivalent to a point but not equivalent to a point but not to the shape consisting of two points

[16]This notion is technically referred to as *homotopy equivalence*, and if the reader has any intuition about this notion, then let us make clear that this is exactly the notion to which we are referring.

$*$ $\simeq$  $\not\simeq$ $* \quad *$

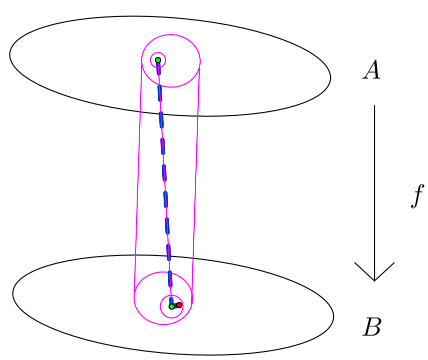
Now, take a map $f : A \to B$ between shapes. For any point b on B we can visualize the shape of all those points in A such that f maps to some point b' to which b is connected by a path. We call this shape the *fiber* of f over b and we can visualize it as the "cylinder shape" traced out in magenta below:



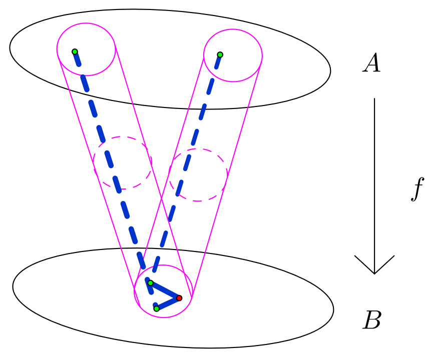
A point of this fiber can be visualized as the pair of blue lines above, i.e. as line (not a path!) connecting x to $f(x)$ together with a path from $f(x)$ to b. Following this visual intuition, the fiber (of map f over a point b) can thus be formalized as the following total shape

$$\mathsf{fib}(f,b) =_{\mathrm{df}} \underset{x\,:\,A}{\mathsf{Total}}\ \mathsf{Path}_B(f(x),b)$$

Now, to say that a fiber is contractible is to say that all these pairs of 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 x to b:



And if this process can be carried out for *every* point of B, viz. if every fiber of f is contractible, then f would intuitively define a deformation of A into B, since it will map, up to a path, every point of A to precisely one point of B. On the other hand, if 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. Spatially, this captures 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.

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. Thus, to observe that a map $A \to B$ is an equivalence is to observe a map that takes every point of B to a center of contraction for its fiber. This allows us to visualize the following formal definition of equivalence

$$\mathsf{isequiv}(f) =_{\mathrm{df}} \underset{x:\,B}{\mathsf{Map}} \ \mathsf{iscontractible}(\mathtt{fib}(f,x))$$

Identity is an Equivalence

The identity map id_A on a shape A can be visualized as the map from A to A which take a point to itself. This identity is an equivalence, since a shape clearly deforms to itself if we just do nothing to it. This is spatially obvious and it visualizes the following rule:

$$\frac{\Gamma \vdash A \textbf{ Shape}}{\Gamma \vdash J : \mathsf{isequiv}(\mathsf{id}_A)} \quad \mathtt{Path\text{-}Elim}$$

Since the statement that the identity map on a shape A is an equivalence is symmetric to the statement that the fiber of every point on that shape is contractible, then we can visualize the center of contraction for these fibers to be exactly the pair given by the point and the trivial path on that point:

$$\frac{\Gamma \vdash A \textbf{ Shape} \qquad \Gamma \vdash a : A}{\Gamma \vdash \mathsf{pr1}(J(a)) \equiv \langle a, \mathsf{refl}_a \rangle} \quad \mathtt{Path\text{-}comp\text{-}1}$$

$$\frac{\Gamma \vdash A\ \textbf{Shape} \qquad \Gamma \vdash a : A}{\Gamma \vdash \mathsf{pr2}(J(a))(\langle a, \mathsf{refl}_a \rangle) \equiv \mathsf{refl}_{\langle a, \mathsf{refl}_a \rangle}}\ \texttt{Path-comp-2}$$

This completes the visualization of the rules for the path shapes. The rules that we have chosen to visualize are not how the rules for "identity types" of MLTT are usually presented, but our rules are in fact equivalent (in the sense of being inter-admissible) to the usual rules for the identity type of MLTT.[17] We chose to present them this way because they seem to us much easier to visualize spatially - but to do so required us to detour through the notion of contractibility and equivalence.[18]

0-shape as the empty shape

There is a shape with no points.

$$\frac{}{\Gamma \vdash \mathbf{0}\ \textbf{Shape}}\ \texttt{0-Form}$$

This is the *empty shape*, which we can visualize as the absence of shape in space. Now assume that we could observe a point e on the empty shape. Such a point e of the empty shape would be a point "in the void" between shapes. And so if from that point you can observe any shape A then there is nothing separating you from "moving closer" to A in order to observe a point on it (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 (a point in the "void") then you are given a foothold to make any observation whatsoever. Therefore, given a place to "stand on" in the void you can move around the void freely, making every observation possible: you can move, as it were, without friction. 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 visualization of *ex falso quodlibet* is that having a point in the void allows you to observe anything whatsoever:

$$\frac{\Gamma \vdash A\ \textbf{Shape} \qquad \Gamma \vdash e : \mathbf{0}}{\Gamma \vdash \texttt{efq}_e : A}\ \texttt{0-Elim}$$

ℕ-shape as the infinite discrete shape

ℕ is to be understood as the infinite discrete shape, i.e. a shape where there are infinitely many points not connected by paths.

Spatially, this means that we can always observe a point on ℕ and if we can observe a point n on ℕ then we can observe another point $s(n)$. In other

[17]Indeed, given the rules for transport above, we can derive the usual J-rule for identity types from the J which witnesses the fact that the identity map is an equivalence.

[18]It must be noted, however, that our presentation relies on the availability of Total and Map constructors, and the equivalence between the two sets of rules holds only modulo their availability.

words, we visualize the shape $\mathbb{N}$ as being "indefinitely extendible".[19] I believe the spatial intuition for such a shape, and the fact that it is reasonable to assume it exists, is clear, and it helps us visualize the following two rules:

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

Now let C be a shape we can observe from the viewpoint of $\mathbb{N}$. Suppose we can observe a point on the portion of C over 0 - which ensures that the total shape of all portions of C above $\mathbb{N}$ is non-empty. Now assume that from the viewpoint consisting of any point z of the total shape $\underset{x\colon \mathbb{N}}{\mathsf{Total}}\, C(x)$ we can observe a point c_s on the portion of C over the successor of the first projection of z. That is, for any pair of points x on $\mathbb{N}$ and y on C over x we can observe a point c_s on C over $s(x)$ – namely, we can always "skip ahead" on $\mathbb{N}$, look "up", and still observe a point on (the portion of) C (above from where we are standing). Then, clearly, given any point on $\mathbb{N}$ we can observe a point on C above n in such a way that if n is 0 that point is c_0 and if n is $s(m)$ then that point is c_s. And this is exactly how we can visualize the following rules:

$$\frac{\begin{array}{c}\Gamma, x\colon \mathbb{N} \vdash C \ \textbf{Shape} \qquad \Gamma \vdash c_0\colon C[0/x] \\ \Gamma, z : \underset{x\colon \mathbb{N}}{\mathsf{Total}}\, C(x) \vdash c_s\colon C[s(\mathsf{pr1}(z))/x] \qquad \Gamma \vdash n\colon \mathbb{N}\end{array}}{\Gamma \vdash \mathtt{ind}_{c_0,c_s,n}\colon C[n/x]}\ \mathbb{N}\texttt{-Elim}$$

$$\frac{\begin{array}{c}\Gamma, x\colon \mathbb{N} \vdash C \ \textbf{Shape} \qquad \Gamma \vdash c_0\colon C[0/x] \\ \Gamma, z : \underset{x\colon \mathbb{N}}{\mathsf{Total}}\, C(x) \vdash c_s\colon C[s(\mathsf{pr1}(z))/x]\end{array}}{\Gamma \vdash \mathtt{ind}_{c_0,c_s,0} \equiv c_0\colon C[0/x]}\ \mathbb{N}\texttt{-Comp-1}$$

$$\frac{\begin{array}{c}\Gamma, x\colon \mathbb{N} \vdash C \ \textbf{Shape} \qquad \Gamma \vdash c_0\colon C[0/x] \\ \Gamma, z : \underset{x\colon \mathbb{N}}{\mathsf{Total}}\, C(x) \vdash c_s\colon C[s(\mathsf{pr1}(z))/x] \qquad \Gamma \vdash n\colon \mathbb{N}\end{array}}{\Gamma \vdash \mathtt{ind}_{c_0,c_s,s(n)} \equiv c_s\colon C[s(n)/x]}\ \mathbb{N}\texttt{-Comp-2}$$

[19] 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 [16]. 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.

$\mathcal{U}_i$-shapes as shapes of shapes

We can imagine that space itself is a shape (living in some "higher" space). Then we have the shape of all shapes: a *universe shape*. In fact, we can visualize a hierarchy of such universes $\mathcal{U}_i$ for each $i \in \omega$

$$\frac{i \in \omega}{\Gamma \vdash \mathcal{U}_i \ \textbf{Shape}} \quad \mathcal{U}_i\texttt{-Form}$$

More precisely, each $\mathcal{U}_i$ is a "shape whose points are shapes". Every point on a universe is a shape and if we can observe it as a point then we can observe it as a shape, which visualizes the following rule:

$$\frac{\Gamma \vdash a : \mathcal{U}_i}{\Gamma \vdash \mathsf{Pts}(a) \ \textbf{Shape}} \quad \mathcal{U}\texttt{-Points}_i$$

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. Indeed, this is exactly how we to visualize Pts in the rule $\mathcal{U}$-$\texttt{Points}$: as the operation of "blowing up" points of a universe into actual shapes. This visual also justifies all the "closure" rules ($\mathcal{U}_i$-Map-1, $\mathcal{U}_i$-Map-2, $\mathcal{U}_i$-Total-1 etc.) given in the Appendix, which we will not write out in full here. Each such rule states that a point on a universe constructed by the shape-constructors we have so far seen is symmetric, and therefore identical to, its "underlying" shape. This is intuitively immediate since the meaning of $\mathcal{U}_i$ is fixed as exactly the shape whose points are shapes.

Every universe $\mathcal{U}_i$ is itself a shape and therefore can itself be observed as a point u_i in a shape of shapes, i.e. in a universe that contains it. One can think of a universe shape as shrinking down to a point, and being observed as a point from the "next higher" universe, and this spatial picture is how we visualize the following rules (for any $i \in \omega$):

$$\frac{}{\Gamma \vdash u_i : \mathcal{U}_{i+1}} \quad \mathcal{U}\texttt{-Accum-1}_i \qquad\qquad \frac{}{\Gamma \vdash \mathsf{Pts}(u_i) \equiv \mathcal{U}_i \ \textbf{Shape}} \quad \mathcal{U}\texttt{-Accum-1}_i$$

Finally, every point of a universe (i.e. every shape in that universe) is a point of every other universe which contains it: shapes, as points, accumulate. The way to visualize this is that if we can observe a shape, then we can observe that shape as a point of a universe. But every universe which contains this universe (as a point) must also contain exactly the shape that we have just observed – and this is visualized by the following rule:

$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad i < j}{\Gamma \vdash a : \mathcal{U}_j} \quad \mathcal{U}\texttt{-cumul}_{i,j}$$

Propositional Resizing for $\mathcal{U}_i$

Some shapes, as we have said, are *contractible*, i.e. equivalent to a point. There is another class of shapes that have contractible path shapes, in the sense that for any two points, *if they exist*, the shape of paths between them is contractible. Such shapes in UF are called *propositions*. Propositions are so called because, up to equivalence, they are truth values: as shapes they are either empty ("false") or contractible ("true").

It is then reasonable to think of "propositions" as shapes such that for any two of their points (which may not exist) there is a map that picks out a path that connects these two points. Given the meaning of the formal symbols introduced so far, this description can be made formal in order to define the property that the shape A "is a proposition" as

$$\mathsf{isaprop}(A) =_{\mathrm{df}} \underset{x,y:A}{\mathsf{Map}}\ \mathsf{Path}_A(x,y)$$

Shapes that are propositions in this sense can be visualized as being very small. Indeed, if they are non-empty, they will be as "small" as a point. As such, we can imagine that if a shape is a proposition in some universe, then it is also a shape in any other universe – and in particular in the "lowest" universe. As it were, such shapes are small enough to thread themselves up and down the universe hierarchy, so to speak. In particular, propositions can be proven, and held to be true, in any universe, and we can always "pull them down" to the lowest universe. And this is exactly how we visualize the following rule:

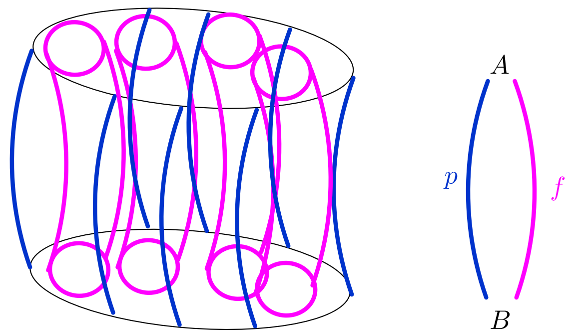
$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad \Gamma \vdash t : \mathsf{isaprop}(a)}{\Gamma \vdash a : \mathcal{U}_0} \quad \texttt{Propositional Resizing}_i$$

Univalence for $\mathcal{U}_i$

As we have seen, an equivalence between shapes A and B is to be visualized as a continuous deformation that transforms A to B without puncturing or tearing them. As we have said, an intuitive way of visualizing this notion is by picturing a map $f : A \to B$ that establishes a "one-to-one correspondence" *up to a path*. Building on what we have said so far the *shape of equivalences* (or *equivalence shape*) is captured by the following definition:

$$A \simeq B =_{\mathrm{df}} \underset{f:\, A \to B}{\mathsf{Total}} \mathsf{isequiv}(f)$$

Now, with these visualizations in mind, we can ask: if two shapes are regarded as points in a universe, then how should we visualize a path between them (i.e. how should we visualize the inhabitants of the shape $\mathsf{Path}_{\mathcal{U}}(A,B)$)? These paths can be visualized as "tubes" connecting the "blown up" shapes one to the other. Clearly, paths between shapes-as-points-on-a-universe and equivalences between shapes are closely related. 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.

Now, each such p can be deformed into f and vice versa. We visualize these deformations as follows: each blue tube can be uniquely broken down into the magenta cylinders induced by an equivalence; conversely the cylinders of an equivalence can each be be uniquely merged into the tube induced by a path. And these processes are visualized exactly as the fact that the shape $A \simeq B$ of equivalences between A and B can be deformed into the shape of paths between them (in $\mathcal{U}$) – which means, precisely, that these two shapes are themselves equivalent. In other words, equivalences are equivalent to paths – which is precisely how we visualize the axiom of univalence (for a given universe $\mathcal{U}_i$):

$$\frac{\Gamma \vdash A : \mathcal{U}_i \qquad \Gamma \vdash B : \mathcal{U}_i}{\Gamma \vdash \mathsf{univ}^i_{A,B} : (A \simeq B) \simeq (\mathsf{Path}_{\mathcal{U}_i}(A, B))} \quad \texttt{Univalence-}i$$

3 Conclusion

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" (or "intuitive"). 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 themselves 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 description of the realm of shapes in Section 2 is 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, and that is certainly a good reason to believe in the consistency of the notion. 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").[20]

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". 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. One can perhaps think of the realm we are describing in the following way. The intension of a shape is a topological space, i.e. a space as usually understood. Its extension, however, is a homotopy type of topological spaces. Namely, the class of topological spaces up to homotopy equivalence.

On the plus side, I believe the spatial meaning explanation in its current form achieves independence from set theory at an intuitive level. At no point during the meaning explanation have I employed talk of elements, collections etc. – namely, at no point have I appealed to some intuitive theory of sets. Thus, insofar as the spatial meaning explanation can be extended to other, stronger HoTTs, we have good reason to claim that UF can achieve *complete* autonomy from set theory, even at the level of intuitive justifications. This would be the first step towards making a convincing case that UF can be a self-standing foundation of mathematics that could plausibly serve as a benchmark of rigour on its very own.

Finally, there may yet be entirely different approaches to providing pre-formal meaning explanations to formalizations of the Univalent Foundations. On the one hand, we could imagine maybe a *non-spatial* pre-formal 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

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

spatial notions I have used in this paper with their "groupoidal" counterparts: for example, to replace "shape" by "structure", "path" by "isomorphism" etc.[21]

To conclude: in this paper I have provided a spatial meaning explanation for HoTT. 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.[22] Exploring the limitations and implications of this new picture presents an urgent task.

Appendix: The Rules of HoTT

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

Structural Rules

$$\frac{}{\bullet\ \textbf{view}}\ \bullet\text{ - }\textbf{view} \qquad \frac{\Gamma\ \textbf{view} \qquad \Gamma \vdash A\ \textbf{Shape}}{\Gamma, x : A\ \textbf{view}}\ \mathsf{ext}\text{ - }\textbf{view}$$

$$\frac{\Gamma, x : A, \Delta\ \ \textbf{view}}{\Gamma, x : A, \Delta \vdash x : A}\ \mathsf{ax}\text{ - }\textbf{view}$$

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

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

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

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

[21] This idea is developed further in [33] based on the formal system in [30].
[22] This has recently been argued in [11].

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

Rules for shape constructors

Total**-shapes**

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, x\colon A \vdash B(x) \ \textbf{Shape}}{\Gamma \vdash \underset{x\colon A}{\mathsf{Total}}\ B(x) \ \textbf{Shape}}\ \mathsf{Total}\text{-}\texttt{Form}$$

$$\frac{\begin{array}{c}\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, x\colon A \vdash B(x) \ \textbf{Shape} \\ \Gamma \vdash a\colon A \qquad \Gamma \vdash b\colon B[a/x]\end{array}}{\Gamma \vdash \langle a, b\rangle\colon \underset{x\colon A}{\mathsf{Total}}\ B(x)}\ \mathsf{Total}\text{-}\texttt{Intro}$$

$$\frac{\Gamma \vdash z\colon \underset{x\colon A}{\mathsf{Total}}\ B(x)}{\Gamma \vdash \mathsf{pr1}(z) : A}\ \mathsf{Total\text{-}pr1} \qquad \frac{\Gamma \vdash z\colon \underset{x\colon A}{\mathsf{Total}}\ B(x)}{\Gamma \vdash \mathsf{pr2}(z) : B[\mathsf{pr1}(z)/x]}\ \mathsf{Total\text{-}pr2}$$

$$\frac{\Gamma \vdash a\colon A \qquad \Gamma \vdash b\colon B[a/x]}{\Gamma \vdash \mathsf{pr1}(\langle a, b\rangle) \equiv a : A \qquad \Gamma \vdash \mathsf{pr2}(\langle a, b\rangle) \equiv b : B[a/x]}\ \mathsf{Total\text{-}comp\text{-}1}$$

$$\frac{\Gamma \vdash z\colon \underset{x\colon A}{\mathsf{Total}}\ B(x)}{\Gamma \vdash z \equiv \langle \mathsf{pr1}(z), \mathsf{pr2}(z)\rangle : \underset{x\colon A}{\mathsf{Total}}\ B(x)}\ \mathsf{Total\ \text{-}comp\text{-}2}$$

Map**-shapes**

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, x\colon A \vdash B(x) \ \textbf{Shape}}{\Gamma \vdash \underset{x\colon A}{\mathsf{Map}}\ B(x) \ \textbf{Shape}}\ \mathsf{Map}\text{-}\texttt{Form}$$

$$\frac{\Gamma, x\colon A \vdash b(x)\colon B(x)}{\Gamma \vdash \lambda(x\colon A).b(x)\colon \underset{x\colon A}{\mathsf{Map}}\ B(x)}\ \mathsf{Map}\text{-}\texttt{Intro}$$

$$\frac{\Gamma \vdash f\colon \underset{x\colon A}{\mathsf{Map}}\ B(x) \qquad \Gamma \vdash a\colon A}{\Gamma \vdash f(a)\colon B[a/x]}\ \mathsf{Map}\text{-}\texttt{Elim}$$

$$\frac{\Gamma, x\colon A \vdash b\colon B(x) \qquad \Gamma \vdash a\colon A}{\Gamma \vdash (\lambda(x\colon A).b)(a) \equiv b[a/x]\colon B[a/x]}\ \mathsf{Map\text{-}Comp}$$

Path-**shapes**

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

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

$$\frac{\begin{array}{c}\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, x : A \vdash B \ \textbf{Shape} \\ \Gamma \vdash a : A \qquad \Gamma \vdash a' : A \\ \Gamma \vdash p : \mathsf{Path}_A(a,a')\end{array}}{\Gamma \vdash t_p\colon B[a/x] \to B[a'/x]} \ \texttt{Path-transport}$$

$$\frac{\begin{array}{c}\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma, x : A \vdash B \ \textbf{Shape} \\ \Gamma \vdash a : A \qquad \Gamma \vdash b : B[a/x]\end{array}}{\Gamma \vdash t_{\mathsf{refl}_a}(b) \equiv b : B[a/x]} \ \texttt{Path-transport-sym}$$

$$\frac{\Gamma \vdash A \ \textbf{Shape}}{\Gamma \vdash J : \mathsf{isequiv}(\mathsf{id}_A)} \ \texttt{Path-Elim}$$

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash a : A}{\Gamma \vdash \mathsf{pr1}(J(a)) \equiv \langle a, \mathsf{refl}_a \rangle} \ \texttt{Path-comp-1}$$

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash a : A}{\Gamma \vdash \mathsf{pr2}(J(a))(\langle a, \mathsf{refl}_a \rangle) \equiv \mathsf{refl}_{\langle a, \mathsf{refl}_a \rangle}} \ \texttt{Path-comp-2}$$

0-shape

$$\frac{}{\Gamma \vdash \mathbf{0} \ \textbf{Shape}} \ \texttt{0-Form} \qquad \frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash e : \mathbf{0}}{\Gamma \vdash \mathsf{efq}_e\colon A} \ \texttt{0-Elim}$$

1-shape

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

$$\frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash a\colon A}{\Gamma \vdash !_a\colon \mathbf{1} \to A} \qquad \frac{\Gamma \vdash A \ \textbf{Shape} \qquad \Gamma \vdash a\colon A}{\Gamma \vdash !_a(a) \equiv a\colon A}$$

$\mathbb{N}$-shape

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

$$\frac{\begin{array}{c}\Gamma, x \colon \mathbb{N} \vdash C\ \textbf{Shape} \\ \Gamma \vdash c_0 \colon C[0/x] \qquad \Gamma, z : \underset{x \colon \mathbb{N}}{\mathsf{Total}}\ C(x) \vdash c_s \colon C[s(\mathsf{pr1}(z))/x] \qquad \Gamma \vdash n \colon \mathbb{N}\end{array}}{\Gamma \vdash \texttt{ind}_{c_0, c_s, n} \colon C[n/x]}$$

$$\frac{\begin{array}{c}\Gamma, x \colon \mathbb{N} \vdash C\ \textbf{Shape} \qquad \Gamma \vdash c_0 \colon C[0/x] \\ \Gamma, z : \underset{x \colon \mathbb{N}}{\mathsf{Total}}\ C(x) \vdash c_s \colon C[s(\mathsf{pr1}(z))/x]\end{array}}{\Gamma \vdash \texttt{ind}_{c_0, c_s, 0} \equiv c_0 \colon C[0/x]}$$

$$\frac{\begin{array}{c}\Gamma, x \colon \mathbb{N} \vdash C\ \textbf{Shape} \qquad \Gamma \vdash c_0 \colon C[0/x] \\ \Gamma, z : \underset{x \colon \mathbb{N}}{\mathsf{Total}}\ C(x) \vdash c_s \colon C[s(\mathsf{pr1}(z))/x] \qquad \Gamma \vdash n \colon \mathbb{N}\end{array}}{\Gamma \vdash \texttt{ind}_{c_0, c_s, s(n)} \equiv c_s \colon C[s(n)/x]}$$

Rules for the Universes

Every rule below is instantiated for any $i \in \omega$, whenever such an i appears.

$$\frac{}{\vdash \mathcal{U}_i\ \textbf{Shape}}\ \texttt{Uni-Form}_i \qquad \frac{\Gamma \vdash a : \mathcal{U}_i}{\Gamma \vdash \texttt{El}(a)\ \textbf{Shape}}\ \mathcal{U}\texttt{-Points}_i$$

$$\frac{\Gamma \vdash a \colon \mathcal{U}_i}{\Gamma \vdash a : \mathcal{U}_j}\ \mathcal{U}\texttt{-Cumul}_{ij},\ \ i < j$$

$$\frac{}{\Gamma \vdash u_i : \mathcal{U}_{i+1}}\ \mathcal{U}\texttt{-Accum-1}_i$$

$$\frac{}{\Gamma \vdash \mathsf{El}(u_i) \equiv \mathcal{U}_i\ \textbf{Shape}}\ \mathcal{U}\texttt{-Accum-2}_i$$

$$\frac{\Gamma \vdash A \colon \mathcal{U}_i \qquad \Gamma \vdash B \colon \mathcal{U}_i}{\Gamma \vdash \mathsf{univ}^i_{A,B} \colon (A \simeq B) \simeq (\mathsf{Path}_{\mathcal{U}_i}(A, B))}\ \texttt{Univalence}_i$$

$$\frac{\Gamma \vdash a \colon \mathcal{U}_i \qquad \Gamma \vdash t : \mathsf{isaprop}(a)}{\Gamma \vdash a \colon \mathcal{U}_0}\ \texttt{Propositional Resizing}_i$$

$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad \Gamma, x : \mathtt{El}(a) \vdash b(x) : \mathcal{U}_i}{\Gamma \vdash \pi(a, x.b(x)) : \mathcal{U}_i}\ \mathcal{U}_i\text{-}\mathsf{Map}\text{ -1}$$

$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad \Gamma, x : \mathtt{El}(a) \vdash b(x) : \mathcal{U}_i}{\Gamma \vdash \mathtt{El}(\pi(a, x.b(x))) \equiv \underset{x : \mathtt{El}(a)}{\mathsf{Map}}\ \mathtt{El}(b(x))\ \textbf{Shape}}\ \mathcal{U}_i\text{-}\mathsf{Map}\text{ -2}$$

$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad \Gamma, x : \mathtt{El}(a) \vdash b(x) : \mathcal{U}_i}{\Gamma \vdash \sigma(a, x.b(x)) : \mathcal{U}_i}\ \mathcal{U}_i\text{-}\mathsf{Total}\text{ -1}$$

$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad \Gamma, x : \mathtt{El}(a) \vdash b(x) : \mathcal{U}_i}{\Gamma \vdash \mathtt{El}(\sigma(a, x.b(x))) \equiv \underset{x : \mathtt{El}(a)}{\mathsf{Total}}\ \mathtt{El}(b(x))\ \textbf{Shape}}\ \mathcal{U}_i\text{-}\mathsf{Total}\text{ -2}$$

$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad \Gamma \vdash b : \mathtt{El}(a) \qquad \Gamma \vdash c : \mathtt{El}(a)}{\Gamma \vdash \mathtt{id}_a(b, c) : \mathcal{U}_i}\ \mathcal{U}_i\text{-}\mathsf{Path}\text{-}1$$

$$\frac{\Gamma \vdash a : \mathcal{U}_i \qquad \Gamma \vdash b : \mathtt{El}(a) \qquad \Gamma \vdash c : \mathtt{El}(a)}{\Gamma \vdash \mathtt{El}(\mathtt{id}_a(b, c)) \equiv \mathsf{Path}_{\mathtt{El}(a)}(b, c)\ \textbf{Shape}}\ \mathcal{U}_i\text{-}\mathsf{Path}\text{-}2$$

$$\frac{}{\vdash \mathtt{z} : \mathcal{U}_i}\ \mathcal{U}_i\text{-}\mathbf{0}\text{-}1 \qquad \frac{}{\vdash \mathtt{El}(\mathtt{z}) \equiv \mathbf{0}\ \textbf{Shape}}\ \mathcal{U}_i\text{-}\mathbf{0}\text{-}2$$

$$\frac{}{\vdash \mathtt{s} : \mathcal{U}_i}\ \mathcal{U}_i\text{-}\mathbf{1}\text{-}1 \qquad \frac{}{\vdash \mathtt{El}(\mathtt{s}) \equiv \mathbf{1}\ \textbf{Shape}}\ \mathcal{U}_i\text{-}\mathbf{1}\text{-}2$$

$$\frac{}{\vdash \mathtt{N} : \mathcal{U}_i}\ \mathcal{U}_i\text{-}\mathbb{N}\text{-}1 \qquad \frac{}{\vdash \mathtt{El}(\mathtt{N}) \equiv \mathbb{N}\ \textbf{Shape}}\ \mathcal{U}_i\text{-}\mathbb{N}\text{-}2$$

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

[14] Gelfand, M.: We do not choose mathematics as our profession, it chooses us: Interview with yuri manin. Notices of the AMS **56**(10) (2009)

[15] Grayson, D.: An Introduction to the Univalent Foundations for Mathematicians. arXiv Preprint (2017). Available at: `https://arxiv.org/pdf/1711.01994.pdf`

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

[17] Hou (Favonia), K.B.: Higher-dimensional types in the mechanization of homotopy theory. Ph.D. thesis, CMU (2017)

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

[30] Tsementzis, D.: First-order logic with isomorphism. arXiv preprint arXiv:1603.03092 (2016)

[31] 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`

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

[33] Tsementzis, D., Halvorson, H.: Foundations and philosophy. Philosopher's Imprint (2017). URL http://philsci-archive.pitt.edu/13504/. Forthcoming

[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)