Max-Planck-Institut für Mathematik
Bonn

A type theory for strictly unital ∞-categories

by

Eric Finster
David Reutter
Jamie Vicary

Max-Planck-Institut für Mathematik
Preprint Series 2020 (60)

Date of submission: November 1, 2020

A type theory for strictly unital ∞-categories

by

Eric Finster
David Reutter
Jamie Vicary

Max-Planck-Institut für Mathematik
Vivatsgasse 7
53111 Bonn
Germany

Computer Laboratory
University of Cambridge
J. J. Thomson Aveneu
Cambridge, CB3 0FD
UK

MPIM 20-60

A Type Theory for Strictly Unital ∞-Categories

Eric Finster*, David Reutter† and Jamie Vicary‡

July 17, 2020

Abstract

We present a type theory for strictly unital ∞-categories, in which a term computes to its strictly unital normal form. Using this as a toy model, we argue that it illustrates important unresolved questions in the foundations of type theory, which we explore. Furthermore, our type theory leads to a new definition of strictly unital ∞-category, which we claim is stronger than any previously described in the literature.

1 Introduction

1.1 Overview

Motivation Standard categorical models of type theory have a curious and well-known feature. When we regard the λ-calculus as a presentation of a cartesian closed category, we must identify all β-equivalent terms, such as the following,

$$(\lambda x.p)q = p[q/x]$$

by quotienting the set of raw syntactic terms by definitional equality, written "=".

However, since any semantic model is required to respect definitional equality, it then becomes impossible for the semantics to represent the dynamics of computation in a nontrivial way. We call this the *Identification Paradox*. Like most good paradoxes, it does not represent a logical absurdity; after all, our standard models of type theory continue to serve us well for many purposes. Rather, it serves to challenge us, as we seek new perspectives on the foundations of computation.

In dependent type theories, the Identification Paradox gains greater weight, since it leads directly to a number well-known meta-theoretic difficulties. The conversion rule

$$\frac{\Gamma \vdash a : A \qquad A = A'}{\Gamma \vdash a : A'}\ \textsc{conv}$$

allows us to silently coerce terms between two types which are equal up to definitional equality, but records no witness of this conversion in the syntax of

*University of Cambridge, UK, ericfinster@gmail.com
†Max Planck Institute for Mathematics, Germany, reutter@mpim-bonn.mpg.de
‡University of Cambridge, UK, jamie.vicary@cl.cam.ac.uk

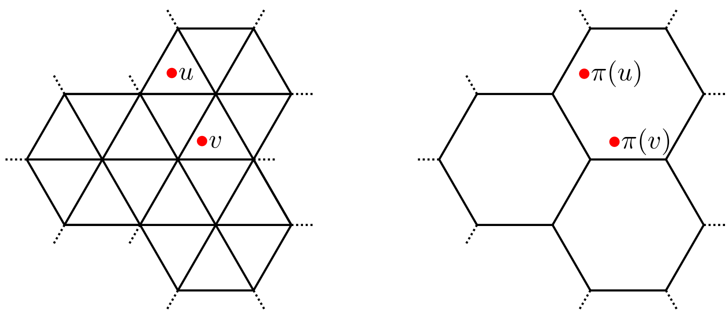
the term a. For typing purposes, definitionally equal types have been *identified*. Such silent conversions can cause serious problems when we attempt to treat the meta-theory of type theory *in* type theory. For example, one extremely natural approach to modeling the syntax of type theory is as a sequence of inductive families: types dependent on contexts, terms dependent on types and so on [10, 11]. However, in this case the "silent" conversion rule is not even type-correct: if we have two syntactically distinct type expressions, say A and A', then our meta-theory forbids the *same* term expression a to have two *distinct* types, whatever our notion of definitional equality may be. Indeed, some kind of explicit coercion becomes necessary because of the proof-relevance of type theory itself. Various partial solutions are possible: working always with raw syntax, using setoids [9], or more recently, using higher inductive types [1]. But it is fair to say that the meta-theory of type theory in a proof-relevant setting hides many subtleties which are not yet fully understood [23]. We regard this situation as symptomatic of the Identification Paradox.

It was suggested already in the 1980s [22] that ideas from *higher category theory* could be relevant to these questions. Put simply, while 1-categorical structures require that we *identify* computationally-equivalent terms in the model, higher categorical structures allow us to retain the information of *why* these terms are equal by modelling them as *higher dimensional* equalities. While various authors have followed this kind of approach in low dimensions [21, 15], many basic foundational questions and difficulties still remain, and the Identification Paradox remains substantially unresolved.

A New Paradigm We now describe a novel approach to resolving the Identification Paradox, illustrated in Figures 1 and 2. We begin by supposing a traditional type theory S as above, and we seek to construct an auxiliary type theory S', with a projection function $\pi : \mathsf{tm}(\mathsf{S}') \to \mathsf{tm}(\mathsf{S})$ on valid syntactic terms that *preserves* definitional equality, but does not *reflect* it, for all $u, v \in \mathsf{tm}(\mathsf{S}')$:

- $u =_{\mathsf{S}'} v \quad \Rightarrow \quad \pi(u) =_{\mathsf{S}} \pi(v)$
- $u =_{\mathsf{S}'} v \quad \not\Leftarrow \quad \pi(u) =_{\mathsf{S}} \pi(v)$

This says intuitively that the auxiliary theory S' is more *fine-grained*, with a weaker notion of definitional equality, conveyed in Figure 1 by the small triangles



(a) The set $\mathsf{tm}(\mathsf{S}')$ with small definitional equality classes

(b) The set $\mathsf{tm}(\mathsf{S})$ with large definitional equality classes

Figure 1: Definitional equality classes of S and S'

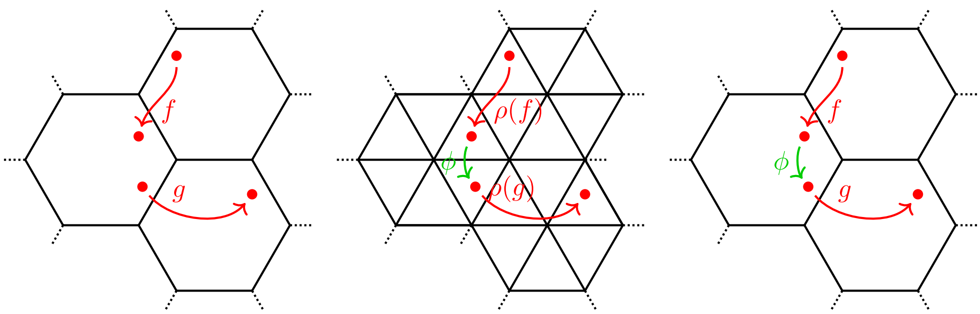
in part (a), compared to the large hexagons in part (b). In particular, we may find some $u, v \in \mathsf{tm}(\mathsf{S}')$ with $u \neq_{S'} v$, but $\pi(u) =_S \pi(v)$, indicated in Figure 1 by the points u, v in different triangles in part (a), and their images $\pi(u), \pi(v)$ in the same hexagon in part (b). Since definitional equality is finer in S', it may therefore have semantics in which u, v are *distinct*, even thought their images in S are definitionally equivalent. Depending on the theory S, this structure (S', π) could perhaps be obtained trivially, possibly by letting the theories have the same terms (with $\pi = \mathrm{id}$), but restricting the notion of definitional equality in S'; that is the idea conveyed in Figure 1, with the triangles and hexagons of parts (a) and (b) tiling the same space, and u, v drawn at the same positions as $\pi(u), \pi(v)$ respectively.

The crux of the idea is then the following. We may also seek a function $\rho : \mathsf{tm}(\mathsf{S}) \to \mathsf{tm}(\mathsf{S}')$ which *preserves and reflects* definitional equality, and which is *cancelled* by π up to definitional equality, for all $p, q \in \mathsf{tm}(\mathsf{S})$:

- $p =_{\mathsf{S}} q \quad \Leftrightarrow \quad \rho(p) =_{\mathsf{S}'} \rho(q)$
- $\pi(\rho(p)) =_{\mathsf{S}} p$

This says intuitively that ρ encodes terms of S as terms of S', in a way which is faithful to their computational interpretation, in such a way that the original definitional equality class can be recovered. It may be surprising that this is possible, since we argued above that S' had a strictly *weaker* notion of definitional equality than S. The conclusion is that the encoding ρ must be nontrivial, allowing the 'broad' definitional equality classes of S to embed in the 'narrow' definitional equality classes of S'.

If we can equip a type theory S with this structure (S', π, ρ), it gives us a way to alleviate the Identification Paradox as follows, which we illustrate in Figure 2. Suppose S is a dependent type theory with types U, V, W, T, with V fibred over T; then for any $t : T$, we have that $V[t]$ is a type. Now suppose we have function types $f : U \to V[t]$ and $g : V[t'] \to W$ in the theory S, with $t =_{\mathsf{S}} t'$. Then the type checker will admit that $V[t] = V[t']$ definitionally, and hence that their composite $f \cdot g$ is a valid term in S by "silent coercion", illustrated in part (a) of the figure; while the target of f does not agree with the source of g, since these points are in the same definitional equality class, the type checker for S will certify the composite as well-typed. We may embed $f \cdot g$ into S' to obtain the valid term $\rho(f \cdot g)$, illustrated in part (b), where we may inspect it with



(a) The term $f \cdot g$ in S (b) The term $\rho(f \cdot g)$ in S' (c) The term $\pi(\rho(f \cdot g))$ in S

Figure 2: Embedding and projecting terms between S and S'

the finer-grained semantics of S' to see in detail the mechanics of the coercion between $V[t]$ and $V[t']$, previously invisible in S, but now manifested as a green arrow ϕ. We may then project it back into S as $\pi(\rho(f \cdot g))$, obtaining a term which is not syntactically identical to the original term $f \cdot g$, but guaranteed to be definitionally equal to it.

One might then ask, if the fine-grained type theory S' is so much more expressive, why should we bother with S at all? The simple answer is that S' may be *much harder* to work with in practice. The smaller definitional equality classes mean that type coercions may not be admitted automatically by the type checker, and must be added 'by hand' when needed. We see this illustrated clearly in Figure 2. In part (a) the terms f, g are composable, since the head of f is in the same definitional equivalence class as the tail of g. But in part (b) they are *not* directly composable, since the head of $\rho(f)$ is in a different definitional equality class to the source of $\rho(g)$; we require the insertion of the explicit coercion ϕ to allow the composite to be formed. While in this instance we constructed ρ algorithmically through the encoding procedure ρ, if we only had access to the fine-grained theory S', we would need to construct it explicitly. The best scenario is to allow ourselves access to *both* S and S', so we can work in S for convenience, then pass to S' when we need more detail about the explicit coercions at play.

Our contribution. We give a comprehensive analysis of a solution to the Identification Paradox along the lines just described, from the perspective of a "toy model", a type theory for weak ∞-categories. At the same time, our results are of independent interest for the theory of weak ∞-categories itself, leading to a new stronger definition of *strictly unital* ∞*-category*[1] than any previously given, and showing the potential impact of type-theoretical approaches to this area of pure mathematics. We summarize our specific contributions as follows.

- For our fine-grained theory S' we choose Catt, a dependent type theory recently introduced [13] to describe weak ∞-categories. Set-theoretic models[2] of Catt correspond to particular weak ∞-categories, and a valid typing judgement $\Gamma \vdash t : A$ can be interpreted as the assertion "in the free weak ∞-category generated by Γ, the term t represents a morphism in the hom-set A".

  The "programs" of this type theory have the curious feature of always being in normal form: they don't "compute" anything at all, and hence the theory has no notion of definitional equality, beyond syntactic equality. The theory is in this sense *maximally* fine-grained. Nonetheless, the theory *describes* something highly nontrivial: the mathematics of weak ∞-categories.

[1] A theory of *weak* ∞*-categories* (also see Section 1.2) is a mathematical formalism for the homotopy theory of composite structures in arbitrary dimensions. A *strictly unital* theory neglects issues related to composing with unit structures; for example, if f is some nontrivial n-dimensional structure, and $\mathbb{1}$ is some n-dimensional unit structure of the appropriate type, then the composite $f \circ \mathbb{1}$ would evaluate to f on-the-nose.

[2] For the type theories we work with in this paper, a *model* is defined to be a functor $\mathsf{C}^{\mathrm{op}} \to \mathsf{Set}$ satisfying a certain gluing condition (see Section 2.3), where C is the category that has contexts Γ of the theory as objects, and definitional equality classes of substitutions $\sigma : \Gamma \to \Delta$ as morphisms.

- For our coarse-grained theory S we choose $\mathsf{Catt}_{\mathsf{su}}$, a dependent type theory for strictly unital ∞-categories, presented here for the first time. As a type theory, it has the interesting property of being identical to Catt, except for the introduction of a nontrivial notion of definitional equality. This type theory yields a new definition of strictly unital ∞-category, as a set-theoretic model of $\mathsf{Catt}_{\mathsf{su}}$. In this theory, definitional equality of terms $t = t'$ can be interpreted as the assertion that the higher-categorical composites represented by t and t' differ only in their unit structure.
- We give a detailed analysis of definitional equality in $\mathsf{Catt}_{\mathsf{su}}$, exhibiting a reduction strategy which produces for every term t a normal form $N(t)$. We show that our reduction strategy generates definitional equality, meaning $t = t'$ if and only if $N(t)$ and $N(t')$ are syntactically identical up to α-equivalence. Definitional equality is hence decidable, and the type theory can be implemented.[3]
- We show that valid $\mathsf{Catt}_{\mathsf{su}}$ terms over disc contexts trivialize, yielding strong evidence that our definition of strictly unital ∞-category is directly comparable, and likely stricter, than previous work (see Related Work section below.)
- The raw syntax of Catt and $\mathsf{Catt}_{\mathsf{su}}$ is identical, and the type checker of $\mathsf{Catt}_{\mathsf{su}}$ is strictly more permissive than that of Catt, so the identity function serves as the projection $\pi : \mathsf{tm}(\mathsf{Catt}) \to \mathsf{tm}(\mathsf{Catt}_{\mathsf{su}})$, which preserves but does not reflect definitional equality, as required.
- By showing in $\mathsf{Catt}_{\mathsf{su}}$ that every term t is *equivalent* to its normalization $N(t)$, we construct an encoding function $\rho : \mathsf{tm}(\mathsf{Catt}_{\mathsf{su}}) \to \mathsf{tm}(\mathsf{Catt})$, satisfying the axioms given on page 3 above. This encoding function algorithmically inserts explicit coercions, which we can inspect directly in Catt, providing a full solution to the Identification Paradox for this toy model.

1.2 Related work

Homotopy Type Theory While our work is not directly concerned with Homotopy Type Theory [24], it is nonetheless heavily inspired by some of the developments that these ideas have provoked in the type theory community. Indeed, the definition of ∞-category presented in [13], and developed in the present work, was based on a similar definition of ∞-groupoid [8]. And this definition in turn can be seen as a distillation of exactly that part of Martin-Löf's identity elimination principle which causes types to behave as higher-dimensional groupoids.

A second point of contact between our work and Homotopy Type Theory is with respect to the problem of defining algebraic structures in a proof-relevant setting. The subtleties of this question are well known to mathematicians [17, 7]. They arise from the observation that, when equality is rendered proof-relevant, the axioms of an algebraic theory may no longer be regarded as *properties* but

[3]The theory has been implemented in OCaml, and the implementation provides additional features not described here, such as implicit arguments for notational convenience. It is available online at https://github.com/ericfinster/catt.io.

rather themselves constitute additional *structure*. And in order to arrive at theories which are well behaved, we must then impose axioms *on the axioms* (so-called coherence conditions). The end result is that the proof-relevant versions of even the most common algebraic structures (e.g. monoids, groups, rings) become infinite dimensional. As an illustration of the difficulties this poses, it remains an open question in Homotopy Type Theory how to define even the correct proof-relevant notion of *category*. While we do not attack this problem directly in this paper, this work may be regarded as a kind of study in the presentation of higher dimensional structures, and we hope that some of the techniques and ideas may prove useful in the future.

Finally, while importing ideas from homotopy theory has proven fruitful for our understanding of the proof-relevant equality of type theory, we feel that, conversely, ideas from logic and computer science may also lead to progress in thinking about higher dimensional structures, along the lines of the unification of physics, topology and logic envisioned in [3]. Indeed, while mathematicians typically treat higher categorical structures starting from *combinatorial* descriptions of their underlying data (for example, using *simplices*, as in [18]), the emerging connections with logic suggest a possible way of thinking of these structures *syntactically*, an approach closer to that of universal algebra. This paper is an example of this line of thought, where we apply type theoretic ideas to the study of a particular structure, that of ∞-categories themselves.

Higher Category Theory Higher category theory is today a significant part of the landscape of modern mathematics. Definitions of higher categories have a long history and, roughly speaking, come in two flavours: 'homotopical', where composition operations are not explicitly specified and compositions are only required to exist up to a contractible space of choices, and 'algebraic', where all composition operations and higher coherences between these are explicitly specified and given as part of the data of a higher category. A prominent representative of this algebraic style is Leinster's variant [17] of Batanin's original definition [4] of a weak ∞-category as an algebra over a certain globular operad.

Our work builds directly on, and extends, the type theory Catt [13], the models of which yield an algebraic notion of weak ∞-category which agrees with a definition due to Maltsiniotis [19], which itself is a close cousin of Batanin's and Leinster's definitions (see [2] for a direct comparison.)

The type theory $\mathsf{Catt}_{\mathsf{su}}$ considered in this paper is a quotient of the type theory Catt by a definitional equality relation, yielding a new notion of strictly unital weak ∞-category. A related definition of strictly unital weak ∞-category in the setting of Leinster's contractible globular operads has been given by Batanin, Cisinski and Weber [5]. We expect our notion of strict unitality to be *stricter* than this previous definition, for the following reason. Their theory includes a notion of *reduction*, corresponding directly to our Theorem 40; and also a notion of *unit compatibility*, corresponding directly to our PRUNE generator for definitional equality. However, one of our generators of definitional equality, which we call ENDO, has no obvious counterpart in their work, and we give on page 18 an example of a pair of terms which are definitionally equal in $\mathsf{Catt}_{\mathsf{su}}$, but we believe would not be identified in BCW's theory.

A prominent difficulty with traditional algebraic definitions of higher category arises from their complexity, with nontrivial computations requiring one to keep track of an abundance of higher-dimensional coherence data. However, a major open conjecture in the field is that much of this data is in fact *redundant*, leading to a search for "semi-strict" replacement theories, in which as much as possible of this coherence data is removed.[4] Low-dimensional examples of such semi-strict theories are well known, with the theory of strict 2-categories (which are equivalent to weak 2-categories) and the theory of Gray-categories [14] (which are equivalent to weak 3-categories) thoroughly understood. A recent framework for a combinatorial theory of semi-strict higher categories, suitable for computer implementation, is described in [12, 20]; however, a disadvantage of that work is that its formal foundations are far from the standard algebraic definitions of Batanin, Leinster, Maltsiniotis and others. Our work seeks ultimately to bridge the gap between traditional algebraic definitions and these more modern combinatorial theories that are more useful in practice, and shows for the first time that type theory can be a powerful tool in this quest.

1.3 Outline

The paper is structured as follows. In Section 2, we set up our type theory for ∞-categories, starting with a presentation along the lines of [13], and then introducing a non-trivial definitional equality. Section 3 analyzes the resulting reduction relation and describes an algorithm for producing normal forms, proving these decide definitional equality. Section 4 describes a certain meta-theoretic property of the resulting system which shows its compatibility with other notions of strictly unital higher category. Finally, in Section 5 we describe our process of "rehydration": starting from a term in our strictly unital theory, we lift this term to a term in the fully weak theory which normalizes to it.

Proofs All proofs are given in the anonymous supplementary materials for the results we present.

2 The Type Theories Catt and $\mathsf{Catt}_{\mathsf{su}}$

It will be convenient to construct our theory in three layers. We begin with the raw syntax and basic rules for contexts, types and substitutions, leaving out the term forming rules. Most of this material is standard, and we make sure to point out any idiosyncracies of notation as we proceed. After introducing the notion of *pasting context* we then present the term forming structure of Catt. Finally, we introduce some combinatorial material necessary to describe our equality relation on terms, culminating in the definition of the theory $\mathsf{Catt}_{\mathsf{su}}$.

2.1 The Base Theory Catt

Raw Syntax We fix an infinite set V of variables, and use lowercase Roman $(x, y, \dots)$ and Greek $(\alpha, \beta, \dots)$ letters to refer to its elements. The raw syntax of Catt consists of four syntactic classes: *contexts*, *types*, *terms* and *substitutions* (denoted Ctx, Type, Term and Sub, respectively). These classes are defined by

[4] It is known that *strict* higher categories, where this coherence data is completely removed, are not sufficiently expressive to model arbitrary phenomena in higher category theory. The challenge is therefore to strictify as far as possible without losing expressivity, but no further; hence the concept of semi-strictness.

the rules in Figure 3. Observe that both contexts and substitutions appear in the raw syntax of terms.

We write $\equiv$ for syntactic equality of the various syntactic classes up to α-equivalence.

Free Variables The *free variables* of elements of each syntactic class are defined by induction on the structure as

$$\begin{aligned} \mathsf{FV}(\emptyset) &= \emptyset & \mathsf{FV}(\star) &= \emptyset \\ \mathsf{FV}(\Gamma, x : A) &= \mathsf{FV}(\Gamma) \cup \{x\} & \mathsf{FV}(s \to_A t) &= \mathsf{FV}(A) \cup \mathsf{FV}(s) \cup \mathsf{FV}(t) \\ \mathsf{FV}(x) &= \{x\} \quad \text{for } x \in V & \mathsf{FV}(\langle\rangle) &= \emptyset \\ \mathsf{FV}(\mathsf{coh}\,(\Gamma : A)[\sigma]) &= \mathsf{FV}(\sigma) & \mathsf{FV}(\langle \sigma, x \mapsto t\rangle) &= \mathsf{FV}(\sigma) \cup \mathsf{FV}(t) \end{aligned}$$

Dimension We define the *dimension* of a type by induction:

$$\mathsf{dim}\,\star = -1 \qquad \mathsf{dim}\,(s \to_A t) = \mathsf{dim}\,A + 1$$

We extend this notion to contexts by asserting that the dimension of a context is one more than the maximum of the dimension of the types occurring in that context.

$$\mathsf{dim}\,\emptyset = -1 \qquad \mathsf{dim}\,(\Gamma, x : A) = \max(\mathsf{dim}\,\Gamma, \mathsf{dim}\,A + 1)$$

Identity Substitutions For a context Γ, we write id_Γ for the identity substitution on Γ defined by

$$\mathrm{id}_\emptyset = \langle\rangle \qquad \mathrm{id}_{\Gamma, x:A} = \langle \mathrm{id}_\Gamma, x \mapsto x\rangle$$

Term Substitutions and Compositions Although the terms of Catt are always in normal form, we will need to perform actual substitutions on terms during type-checking. We therefore also define a semantic form of substitution which calculates by induction on the structure of terms. We denote this operations by

$$\frac{}{\emptyset : \mathsf{Ctx}} \qquad \frac{\Gamma : \mathsf{Ctx} \qquad A : \mathsf{Type}}{\Gamma, A : \mathsf{Ctx}}$$

$$\frac{}{\star : \mathsf{Type}} \qquad \frac{A : \mathsf{Type} \qquad s : \mathsf{Term} \qquad t : \mathsf{Term}}{s \to_A t : \mathsf{Type}}$$

$$\frac{v : V}{v : \mathsf{Term}} \qquad \frac{\Gamma : \mathsf{Ctx} \qquad A : \mathsf{Type} \qquad \sigma : \mathsf{Sub}}{\mathsf{coh}\,(\Gamma : A)[\sigma] : \mathsf{Term}}$$

$$\frac{}{\langle\rangle : \mathsf{Sub}} \qquad \frac{\sigma : \mathsf{Sub} \qquad t : \mathsf{Term}}{\langle \sigma, t\rangle : \mathsf{Sub}}$$

Figure 3: Raw syntax

$[\![-]\!]$ in order to distinguish if from the $[-]$ appearing in coherence terms, which is part of the syntactic structure. This operation is defined as follows:

$$\begin{aligned}
\star[\![\sigma]\!] &= \star \\
(s \to_A t)[\![\sigma]\!] &= s[\![\sigma]\!] \to_{A[\sigma]} t[\![\sigma]\!] \\
\mathsf{coh}\,(\Gamma : s \to t)[\tau][\![\sigma]\!] &= \mathsf{coh}\,(\Gamma : s \to t)[\tau \circ \sigma] \\
x[\![\sigma]\!] &= t \qquad \text{if } x \mapsto t \in \sigma
\end{aligned}$$

where the composition of substitutions ($\circ$) is defined mutually recursively as

$$\langle\rangle \circ \sigma = \langle\rangle \qquad\qquad \langle \tau, t\rangle \circ \sigma = \langle \tau \circ \sigma, t[\![\sigma]\!]\rangle$$

With the previous material in place, our basic typing judgements for contexts, types and substitutions are given in Figure 4. The rules are standard for a dependent type theory. Note that our types consist of just a single base type denoted $\star$ and a formation rule analogous to the formation rule for identity types in Martin Löf Type Theory. The fact that this rule captures faithfully the notion of globular set is at the heart of the connection between type theory and higher category theory, and is the basis of this syntactic description of ∞-categories.

Support Given a term $t : \mathsf{Term}$ and a type $A : \mathsf{Type}$, we can define the *support* of t and A to be the union of their free variables. That is

$$\mathsf{supp}(t, A) = \mathsf{FV}(t) \cup \mathsf{FV}(A)$$

In practice, we will only use this definition when we are given a context Γ and we have $\Gamma \vdash t : A$. When Γ and A are clear from the context, we will often simply write $\mathsf{supp}(t)$ and refer to this as the support of t.

Pasting Contexts The terms of Catt are derived from isolating a distinguished subset of contexts which we call *pasting contexts*. A set of rules for exhibiting evidence that a given context is a pasting context was the key innovation of [13]. These rules are presented in Figure 5.

Boundary Variables For each pasting context $\Gamma \vdash_p$, we will define two distinguished subsets of the variables, denoted $\partial^-(\Gamma)$ and $\partial^+(\Gamma)$. First, for a variable $x : A \in \Gamma$, define its dimension to be $\mathsf{dim} A + 1$. Furthermore, let us say that a

$$\frac{}{\emptyset \vdash} \qquad\qquad \frac{\Gamma \vdash \qquad \Gamma \vdash A : \mathsf{Type}}{\Gamma, x : A \vdash}$$

$$\frac{\Gamma \vdash}{\Gamma \vdash \star : \mathsf{Type}} \qquad\qquad \frac{\Gamma \vdash A : \mathsf{Type} \qquad \Gamma \vdash a : A \qquad \Gamma \vdash b : A}{\Gamma \vdash a \to_A b : \mathsf{Type}}$$

$$\frac{\Gamma \vdash}{\Gamma \vdash \langle\rangle : \emptyset} \qquad\qquad \frac{\Gamma \vdash \sigma : \Delta \qquad \Delta \vdash A : \mathsf{Type} \qquad \Gamma \vdash t : A[\sigma]}{\Gamma \vdash \langle \sigma, x \mapsto t\rangle : \Delta, x : A}$$

Figure 4: Basic typing rules

variable is *target-free* if it does *not* occur as the target of any other variables in Γ. Similarly, we have the notion of *source-free*. We now define:

$$\partial^-(\Gamma) := \{x \in \Gamma \mid \mathsf{dim}\, x < \mathsf{dim}\, \Gamma - 1,\ \textit{or},\ \mathsf{dim}\, x = \mathsf{dim}\, \Gamma - 1 \text{ and } x \text{ is target-free}\}$$
$$\partial^+(\Gamma) := \{x \in \Gamma \mid \mathsf{dim}\, x < \mathsf{dim}\, \Gamma - 1,\ \textit{or},\ \mathsf{dim}\, x = \mathsf{dim}\, \Gamma - 1 \text{ and } x \text{ is source-free}\}$$

Terms With these notions in place, the typing rules for terms of Catt are shown in Figure 6. Note that when writing the substitution in a coherence term, we typically omit the angled brackets, writing $\mathsf{coh}\,(\Gamma : s \to_A t)[a, b, c, \dots]$ instead of $\mathsf{coh}\,(\Gamma : s \to_A t)[\langle a, b, c, \dots \rangle]$.

Examples We record here some basic examples of well-typed terms. In terms of the structure of an ∞-category, these terms represent, respectively: unary composition, binary composition, ternary composition, identities of objects, the right unit law, and a "padding" composition we will describe below. Note that we use parentheses for delineating our contexts here as it improves readability.

$$\mathsf{comp}_1 := \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y) : x \to_\star y)[x, y, f]$$

$$\mathsf{comp}_2 := \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y)(z : \star)(g : y \to_\star z) : x \to_\star z)[x, y, f, z, g]$$

$$\begin{aligned}\mathsf{comp}_3 := \mathsf{coh}\,(&(x : \star)(y : \star)(f : x \to_\star y)(z : \star)(g : y \to_\star z)\\ &(w : \star)(h : z \to_\star w) : x \to_\star w)[x, y, f, z, g, w, h]\end{aligned}$$

$$\mathbb{1}_0 := \mathsf{coh}\,((x : \star) : x \to_\star x)[x]$$

$$\mathsf{unit\text{-}r} := \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y) : \mathsf{comp}_2[\![x, y, f, y, \mathbb{1}_0[\![y]\!]]\!] \to_{x \to_\star y} f)[x, y, f]$$

$$\mathsf{comp}_{2,0} := \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y)(z : \star)(g : y \to_\star z)(h : y \to_\star z)(\alpha : g \to_{y \to_\star z} h)$$

$$\frac{}{x : \star \vdash_p x : \star}\ \star \qquad \frac{\Gamma \vdash_p x : \star}{\Gamma \vdash_p}\ \checkmark$$

$$\frac{\Gamma \vdash_p x : A}{\Gamma, y : A, f : x \to_A y \vdash_p f : x \to_A y}\ \Uparrow \qquad \frac{\Gamma \vdash_p f : x \to_A y}{\Gamma \vdash_p y : A}\ \Downarrow$$

Figure 5: Pasting contexts

$$\frac{\Gamma \vdash \qquad (x : A) \in \Gamma}{\Gamma \vdash x : A}$$

$$\frac{\Gamma \vdash_p \qquad \Gamma \vdash s \to_A t \qquad \Delta \vdash \sigma : \Gamma \qquad \mathsf{supp}(s) = \partial^-(\Gamma) \qquad \mathsf{supp}(t) = \partial^+(\Gamma)}{\Delta \vdash \mathsf{coh}\,(\Gamma : s \to_A t)[\sigma] : s[\![\sigma]\!] \to_{A[\![\sigma]\!]} t[\![\sigma]\!]}$$

$$\frac{\Gamma \vdash_p \qquad \Gamma \vdash s \to_A t \qquad \Delta \vdash \sigma : \Gamma \qquad \mathsf{supp}(s) = \mathsf{supp}(t) = \mathsf{FV}(\Gamma)}{\Delta \vdash \mathsf{coh}\,(\Gamma : s \to_A t)[\sigma] : s[\![\sigma]\!] \to_{A[\![\sigma]\!]} t[\![\sigma]\!]}$$

Figure 6: Terms

$$(w : \star)(k : z \to_\star w)$$
$$: \mathsf{comp_3}[\![x, y, f, z, g, w, k]\!] \to_{(x \to_\star w)} \mathsf{comp_3}[\![x, y, f, z, h, w, k]\!])[x, y, f, z, g, h, \alpha, w, k]$$

From the perspective of the theory of ∞-categories, these all correspond to *operations* from low-dimensional higher category theory, such as the theory of bicategories. In that more familiar language, given objects x, y, z, w and 1-morphisms $f : x \to y$, $g : y \to z$, $h : z \to w$, we have the following, where we write $\cdot$ for forward composition of 1-morphisms:

- $\mathsf{comp_2}[\![x, f, y, z, g]\!]$ corresponds to the binary composite $f \cdot g$;
- $\mathsf{comp_3}[\![x, y, f, z, g, w, h]\!]$ corresponds to the unbiased ternary composite $f \cdot g \cdot h$, which is not directly defined in the traditional notion of bicategory;
- $\mathsf{comp_2}[\![x, \mathsf{comp_2}[\![x, f, y, z, g]\!], z, w, h]\!]$ corresponds to the repeated binary composite $(f \cdot g) \cdot h$;
- $\mathbb{1}_0[\![x]\!]$ corresponds to the 1-morphism id_x;
- $\mathsf{unit\text{-}r}[\![f]\!]$ corresponds to the invertible 2-morphism $f \cdot \mathrm{id}_x \Rightarrow f$.

We could similarly write down operations for other familiar operations in the theory of weak ∞-categories, such as associators, interchangers, and so on, in principle in all dimensions. It is in this sense that Catt gives a formal language for weak ∞-categories. More examples are given in [13].

This last example $\mathsf{comp_{2,0}}$ is part of a family of coherences $\mathsf{comp_{d,k}}$ for $k < d$, which will play a role in Section 5. While we will not give a formal definition, these compositions can be described intuitively as follows: they consist of the "unbiased" composite of a d-dimensional disc D with two $(k+1)$-dimensional discs S and T glued to the k-dimensional source and target of D, respectively. We might write this composite intuitively, as follows:

$$S \circ_k D \circ_k T$$

Identity Terms Generalizing the 0-dimensional case above, we can define an identity on cells of arbitrary dimension. To do so, we assume that the set V of variables contains elements d_i and d'_i for $i \in \mathbb{N}$. Now define the *k-disc context* and the *$(k-1)$-sphere type* by mutal induction on k as follows:

$$\mathrm{D}^0 := \emptyset, d_0 : S^{-1} \qquad S^{-1} := \star$$
$$\mathrm{D}^{k+1} := \mathrm{D}^k, (d'_k : S^{k-1}), (d_{k+1} : S^k) \qquad S^k := d_k \to_{S^{k-1}} d'_k$$

Finally, for $k \in \mathbb{N}$, we define the *identity on the k-disc*, written $\mathbb{1}_k$ as

$$\mathbb{1}_k := \mathsf{coh}\,(\mathrm{D}^k : d_k \to_{S^{k-1}} d_k)[\mathrm{id}_{D^k}]$$

One easily checks that this is a valid term in the context D^k.

2.2 The type theory $\mathsf{Catt}_{\mathsf{su}}$

As described in the introduction, the type theory Catt of the previous section contains no non-trivial definitional equalities: while calculation happens during type-checking, all terms themselves are in normal form. In this section, we introduce our equality relation. Its definition will require some combinatorial preparation which we turn to now.

Labeled Dyck Words Observe that each of the rules for pasting contexts in Figure 5 has at most one hypothesis, and consequently, derivations made with these rules are necessarily linear. In fact, complete derivations of the fact that a context is a pasting context can be identified with *Dyck words* [16]. While strictly speaking, a pasting context is defined as a pair of a syntactic context and a derivation of the fact that it is well-formed, this representation contains quite a bit redundancy and can be somewhat awkward to manipulate in practice. For the purposes of presentation, therefore, it is convenient to introduce a somewhat simplified representation of pasting diagrams which we call *labeled Dyck words*.

Labeled Dyck words may be pictured as a set of up and down moves, with each up move labeled by a pair of elements of some arbitrary labeling set L. Concretely, given L, the set $\mathsf{LDyck}\, L\, n$ of *labeled Dyck words of excess* n is defined by the rules given in Figure 8 (in practice, we often omit L from the notation as it will be clear from context). The parameter n records the difference between the number of up moves and down moves. Note that the definition ensures that the excess is always non-negative, so that we always have at least as many up moves as down moves, leading to the "mountain" diagram shown in Figure 7.

When using labeled Dyck words to represent contexts, we take $L = V$, the set of variables. We hope the reader will notice how the rules for labeled Dyck words mirror exactly the derivation rules for pasting contexts of Figure 5. (The additional ✓ rule for pasting contexts forces a complete derivation $\Gamma \vdash_p$ to be of excess 0). As an example, consider the context

$$(x : *)(y : *)(f : x \to_\star y)(z : *)(g : y \to_\star z)$$

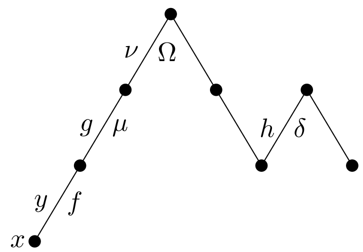


Figure 7: An element of $\mathsf{LDyck}\, L\, 1$
$\Downarrow (\Uparrow (\Downarrow (\Downarrow (\Uparrow (\Uparrow (\Uparrow (\star x)\, y\, f)\, g\, \mu)\, \nu\, \Omega)))\, h\, \delta)$

It is proven to be a pasting context via the derivation

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{(x:\star)\vdash_p (x:\star)}\star}{(x:\star)(y:\star)(f:x\to_\star y)\vdash_p (f:x\to_\star y)}\Uparrow}{(x:\star)(y:\star)(f:x\to_\star y)\vdash_p (y:\star)}\Downarrow}{(x:\star)(y:\star)(f:x\to_\star y)(x:\star)(g:y\to_\star z)\vdash_p (g:y\to_\star z)}\Uparrow}{(x:\star)(y:\star)(f:x\to_\star y)(x:\star)(g:y\to_\star z)\vdash_p (z:\star)}\Downarrow$$

and its labeled Dyck word representation is

$$\Downarrow\ (\Uparrow\ (\Downarrow\ (\Uparrow\ (\star x)\, y\, f))\, z\, g)) : \mathsf{LDyck}\, V\, 0$$

Notice that the only difference between these two representations is that the pasting context, together with its derivation, remembers all the typing information of the all variables, while the Dyck word representation remembers just the variable names. Since the types can be recovered from the structure of the Dyck word itself, this represents no loss of information. In order to fix notation, if $\Gamma \vdash_p$ is a pasting context, we write $\lfloor\Gamma\rfloor : \mathsf{LDyck}\, 0$ for the corresponding labeled Dyck word and, conversely, for a Dyck word $d : \mathsf{LDyck}\, 0$, we write $\lceil d\rceil$ for the corresponding pasting context.

$$\dfrac{x\in L}{\star x : \mathsf{LDyck}\, L\, 0} \qquad \dfrac{n:\mathbb{N} \qquad d:\mathsf{LDyck}\, L\, n \qquad y:L \qquad f:L}{\Uparrow\ d\, y\, f : \mathsf{LDyck}\, L\, (S\, n)}$$

$$\dfrac{n:\mathbb{N} \qquad d:\mathsf{LDyck}\, L\, (S\, n)}{\Downarrow\ d : \mathsf{LDyck}\, L\, n}$$

Figure 8: Labeled Dyck words

A nice advantage of this representation is that we can write programs on pasting contexts using pattern matching style (for which we will use an Agda-style syntax). As an example, consider the function

$$\mathsf{flatten} : \{n:\mathbb{N}\}(d:\mathsf{LDyck}\, n) \to \mathsf{List}\, L$$

$$\dfrac{n:\mathbb{N} \qquad d:\mathsf{LDyck}\, n \qquad y:L \qquad f:L}{\Updownarrow_{\mathsf{pk}}\ d : \mathsf{Peak}\, (\Downarrow\, (\Uparrow\ d\, y\, f))}$$

$$\dfrac{n:\mathbb{N} \qquad d:\mathsf{LDyck}\, n \qquad y:L \qquad f:L \qquad p:\mathsf{Peak}\, d}{\Uparrow_{\mathsf{pk}}\ d\, y\, f\, p : \mathsf{Peak}\, (\Uparrow\ d\, y\, f)}$$

$$\dfrac{n:\mathbb{N} \qquad d:\mathsf{LDyck}\, (S\, n) \qquad p:\mathsf{Peak}\, d}{\Downarrow_{\mathsf{pk}}\ d\, p : \mathsf{Peak}\, (\Downarrow\ d)}$$

Figure 9: Peaks

extracting the list of labels. This function can be defined as

$$\begin{aligned}
&\mathsf{flatten}\,(\star x) = [x] \\
&\mathsf{flatten}\,(\Uparrow\ d\,y\,f) = [\mathsf{flatten}\,d, y, f] \\
&\mathsf{flatten}\,(\Downarrow\ d) = \mathsf{flatten}\,d
\end{aligned}$$

We will make use of this function below.

Peaks A special role will be played in what follows by the positions in a label Dyck word where we change direction from moving up to moving down. We call these the *peaks*. It is not hard to give an induction characterization of peaks, and we do so in Figure 9.

We can use the definition of peaks to manipulate our Dyck word in various ways. Two operations which will be useful in what follows are *excising* a peak, and *replacing* the labels which occur at a given peak. The first operation has type

$$\mathsf{excise} : \{n : \mathbb{N}\}(d : \mathsf{LDyck}\,n)(p : \mathsf{Peak}\,d) \to \mathsf{LDyck}\,n$$

and can be defined as

$$\begin{aligned}
&\mathsf{excise}\,_{_}\,(\Updownarrow_{pk}\ d\,y\,f) = d \\
&\mathsf{excise}\,_{_}\,(\Uparrow_{pk}\ d\,y\,f\,p) = \Uparrow (\mathsf{excise}\,p)\,y\,f \\
&\mathsf{excise}\,_{_}\,(\Downarrow_{pk}\ d\,p) = \Downarrow (\mathsf{excise}\,p)
\end{aligned}$$

The second operation is similar taking as arguments the new labels for the specified peak:

$$\mathsf{replace} : \{n : \mathbb{N}\}(d : \mathsf{LDyck}\,n)(p : \mathsf{Peak}\,d)(g : L)(\alpha : L) \to \mathsf{LDyck}\,n$$

It is defined as

$$\begin{aligned}
&\mathsf{replace}\,_{_}\,(\Updownarrow_{pk}\ d\,y\,f)\,g\,\alpha = \Downarrow (\Uparrow\ d\,g\,\alpha) \\
&\mathsf{replace}\,_{_}\,(\Uparrow_{pk}\ d\,y\,f\,p)\,g\,\alpha = \Uparrow (\mathsf{replace}\,p\,g\,\alpha)\,y\,f \\
&\mathsf{replace}\,_{_}\,(\Downarrow_{pk}\ d\,p)\,g\,\alpha = \Downarrow (\mathsf{replace}\,p\,g\alpha)
\end{aligned}$$

A pictorial representation of the excision operation is given in Figure 10.

Locally Maximal Variables If $\Gamma \vdash_p$ is a pasting context, we say that a variable which occurs as the label of a peak in the Dyck word representation of Γ is *locally maximal.* Intuitively speaking, such variables represent those cells which are of highest dimension "in their neighborhood." We write $\mathsf{LM}(\Gamma)$ for the set of locally maximal variables and for $\alpha \in \mathsf{LM}(\Gamma)$ we write p_α for the corresponding peak. Observe that if $\alpha \in \mathsf{LM}(\Gamma)$, we may speak of the *dimension* $\mathsf{dim}\,\alpha$ since

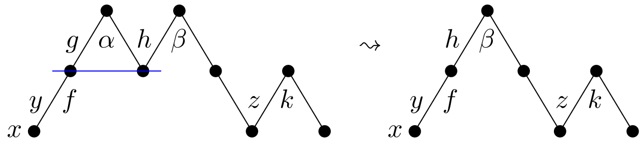


Figure 10: Excising a peak

α is assigned a type by Γ. Moreover, since all locally maximal variables are positive dimensional, any locally maximal variable $\alpha \in \mathsf{LM}(\Gamma)$ must be assigned an arrow type $\alpha : s \to_A t$ for some type A. In such a situation, we refer to s as the *source* of α and to t as its *target*, denoted $\mathsf{src}(\alpha)$ and $\mathsf{tgt}(\alpha)$ respectively.

Given a pasting context $\Gamma \vdash_p$ and a locally maximal variable $\alpha : \mathsf{LM}(\Gamma)$, we write $\Gamma /\!/ \alpha$ for the pasting context obtained by excising the peak where α occurs in the Dyck word representation of Γ.

$$\Gamma /\!/ \alpha := \lceil \mathsf{excise}\ \lfloor \Gamma \rfloor\ p_\alpha \rceil$$

There is a natural substitution $\Gamma /\!/ \alpha \vdash \pi_\alpha : \Gamma$ which may be described as follows: regard $\lfloor \Gamma \rfloor$ as a Dyck word labeled by *terms* as opposed to merely variables. Then we can extract a substitution by replacing the label at the peak p_α with an identity term as follows:

$$\pi_\alpha := \lceil \mathsf{flatten}\ (\mathsf{replace}\ \lfloor \Gamma \rfloor\ p_\alpha\ (\mathsf{src}(\alpha))\ \mathbb{1}_{\mathsf{dim}\alpha}) \rceil$$

Finally, if Δ is another context, and we are given a substitution $\Delta \vdash \sigma : \Gamma$, then we have a substitution $\Delta \vdash \sigma /\!/ \alpha : \Gamma /\!/ \alpha$ obtained by merely dropping the terms assigned to the variables α and $\mathsf{tgt}(\alpha)$,which are necessarily adjacent, from σ:

$$\sigma := [\ldots, \mathsf{src}(\alpha), \mathsf{tgt}(\alpha), \alpha, \ldots] \quad \Rightarrow \quad \sigma /\!/ \alpha := [\ldots, \mathsf{src}(\alpha), \ldots]$$

In this case, the elements t, α are removed from the list to obtain the quotient substitution.

Unfolding Types Given a type T, we obtain a list of terms $\{T\}$ as follows:

$$\{\star\} := [] \qquad\qquad \{u \to_A v\} := [\{A\}, u, v]$$

This is useful to define our endo-coherence removal relation below. We now prove a small lemma regarding type unfolding.

Lemma 1. *Substitution is compatible with type unfolding:*

$$\{T[\![\sigma]\!]\} \equiv \{T\} \circ \sigma$$

Proof. If $T \equiv \star$ then the lemma is vacuously true, since $\{\star\}$ is the empty list. If $T \equiv u \to_A v$, then we reason by induction on subtypes as follows:

$$\{T[\![\sigma]\!]\} \equiv \{(u \to_A v)[\![\sigma]\!]\} \equiv [\{A[\![\sigma]\!]\}, u[\![\sigma]\!], v[\![\sigma]\!]] \equiv [\{A\}, u, v] \circ \sigma \equiv \{u \to_A v\} \circ \sigma$$

This completes the proof. □

$$\frac{\Gamma \vdash_p \qquad A : \mathsf{Type} \qquad \sigma : \mathsf{Sub} \qquad \tau : \mathsf{Sub} \qquad \alpha : \mathsf{LM}(\Gamma) \qquad \alpha[\sigma] = \mathbb{1}_{\mathsf{dim}\alpha - 1}[\tau]}{\mathsf{coh}\ (\Gamma : A)[\sigma] = \mathsf{coh}\ (\Gamma /\!/ \alpha : A[\pi_\alpha])[\sigma /\!/ x]}\ \textsc{prune}$$

$$\frac{n : \mathbb{N} \qquad \sigma : \mathsf{Sub}}{\mathsf{coh}\ (\mathrm{D}^{n+1} : S^n)[\sigma] = d_{n+1}[\sigma]}\ \textsc{disc}$$

$$\frac{\Gamma : \mathsf{Ctx} \qquad A : \mathsf{Type} \qquad t : \mathsf{Term} \qquad u : \mathsf{Term} \qquad t = u \qquad \sigma : \mathsf{Sub}}{\mathsf{coh}\ (\Gamma : t \to_A u)[\sigma] = \mathbb{1}_{\mathsf{dim}A+1}[\{A[\sigma]\}, t[\sigma]]}\ \textsc{endo}$$

Figure 11: Generating equality judgements on terms

Equality in $\mathsf{Catt_{su}}$ With these definitions in place, we can now define our equality relation on terms in Figure 11. There are three generating equalities, which we summarize as follows.

- *Pruning* scans the locally maximal arguments of a substitution looking for identity terms. When such an argument appears, it may be removed while at the same time removing the corresponding locally maximal cell from the pasting diagram defining the coherence.
- *Disc Removal* asserts that unary composites may be removed from the head of a term.
- *Endomorphism Coherence Removal* asserts that coherences associated to a repeated term may be replaced with identities on that term.

From here, we extend this equality relation to contexts, types and substitutions by structural induction on the formations rules, and close these relations under reflexivity, symmetry and transitivity. As an example, the equality relation on types are shown in Figure 12. Contexts, terms and substitutions are similar and we omit the details.

$$\frac{A : \mathsf{Type}}{A = A} \qquad \frac{A : \mathsf{Type} \quad A' : \mathsf{Type} \quad A = A'}{A' = A}$$

$$\frac{A : \mathsf{Type} \quad A' : \mathsf{Type} \quad A'' : \mathsf{Type} \quad A = A' \quad A' = A''}{A = A''}$$

$$\frac{A : \mathsf{Type} \quad A' : \mathsf{Type} \quad A = A' \quad s : \mathsf{Term} \quad t : \mathsf{Term}}{s \to_A t = s \to_{A'} t}$$

$$\frac{A : \mathsf{Type} \quad s : \mathsf{Term} \quad s' : \mathsf{Term} \quad s = s' \quad t : \mathsf{Term}}{s \to_A t = s' \to_A t}$$

$$\frac{A : \mathsf{Type} \quad s : \mathsf{Term} \quad t : \mathsf{Term} \quad t' : \mathsf{Term} \quad t = t'}{s \to_A t = s \to_A t'}$$

Figure 12: Equality relation on types

Finally, to integrate our new equality with the type system, we require all typing judgements to be equipped with conversion rules. These rules are listed in Figure 13.

Example Reductions We record here examples of our three generating reductions, in order to give the reader a flavour of how they operate.

- *Pruning* The pruning relation is the workhorse of the theory, leading in practice to the richest algebraic phenomena. This relation says the following: if a coherence term has a substitution which sends a locally-maximal variable α to an identity, then α can be removed completely from the corresponding pasting context. Let us examine the behavior of this relation on the following term

$$\mathsf{comp}_2[\![x, y, f, y, \mathbb{1}_0[\![y]\!]]\!] := \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y)$$

$$(z:\star)(g:y\to_\star z):x\to_\star z)[x,y,f,y,\mathbb{1}_0[\![y]\!]]$$

Observe that g occurs in a locally maximal position in the head context:

$$\Gamma := (x:\star)(y:\star)(f:x\to_\star y)(z:\star)(g:y\to_\star z)$$

Moreover, the argument supplied in this position is $\mathbb{1}_0[\![y]\!]$. Hence the pruning relation applies. We have

$$\begin{aligned}\Gamma /\!\!/ g &:= (x:\star)(y:\star)(f:x\to_\star y)\\ \pi_g &:= [x,y,f,y,\mathbb{1}_0[\![y]\!]]\\ [x,y,f,y,\mathbb{1}_0[\![y]\!]] /\!\!/ g &:= [x,y,f]\end{aligned}$$

so that this term is definitionally equal to the following:

$$\mathsf{coh}\,((x:\star)(y:\star)(f:x\to_\star y):x\to_\star y)[x,y,f]$$

• *Disc Removal* This relation can be summarized as follows: when the head of a coherence term has a certain form, the entire term is definitionally equal to the last argument of its substitution.

The result of the previous example may have been somewhat surprising: it would have been natural to expect that composing the arrow f with an identity reduced to f itself, whereas we have instead been left with a non-trivial coherence term. Upon closer inspection, we see that the resulting term can be considered the "unary composite" of f. The existence of these composites arise from the fact that our definition of ∞-category is *unbiased*: it allows (vertical) n-ary composites of cells for all n, including $n = 1$. Our second definitional equality exactly removes these superfluous unary composites. In the case at hand, application of the Disc rule now yields:

$$\mathsf{coh}\,((x:\star)(y:\star)(f:x\to_\star y):x\to_\star y)[x,y,f] = f$$

$$\frac{\Gamma:\mathsf{Ctx}\qquad \Gamma':\mathsf{Ctx}\qquad \Gamma=\Gamma'\qquad \Gamma\vdash}{\Gamma'\vdash}$$

$$\frac{\Gamma:\mathsf{Ctx}\qquad \Gamma':\mathsf{Ctx}\qquad \Gamma=\Gamma'\qquad A:\mathsf{Type}\qquad A':\mathsf{Type}\qquad A=A'\qquad \Gamma\vdash A}{\Gamma'\vdash A'}$$

$$\frac{A:\mathsf{Type}\qquad A':\mathsf{Type}\qquad \begin{matrix}\Gamma:\mathsf{Ctx} & \Gamma':\mathsf{Ctx} & \Gamma=\Gamma'\\ A=A' & t:\mathsf{Term} & t':\mathsf{Term}\end{matrix}\qquad t=t'\qquad \Gamma\vdash t:A}{\Gamma'\vdash t':A'}$$

$$\frac{\Delta:\mathsf{Ctx}\qquad \Delta':\mathsf{Ctx}\qquad \begin{matrix}\Gamma:\mathsf{Ctx} & \Gamma':\mathsf{Ctx} & \Gamma=\Gamma'\\ \Delta=\Delta' & \sigma:\mathsf{Sub} & \sigma':\mathsf{Sub}\end{matrix}\qquad \sigma=\sigma'\qquad \Gamma\vdash\sigma:\Delta}{\Gamma'\vdash\sigma':\Delta'}$$

Figure 13: Conversion rules

Note that this replaces the coherence term with the final argument, f, of its substitution $[x, y, f]$. As f is now a variable, the term is now in normal form (see Section 3).

• *Endomorphism Coherence Removal* Another curious redundancy of the fully weak definition of ∞-category is the existence of "fake identities": cells which are "morally" the identity on some composite cell, but do not have an identity coherence at their head. As an example, consider the term:

$$\begin{aligned}&\mathsf{coh}\,((x:\star)(y:\star)(f:x\to_\star y)(z:\star)(g:y\to_\star z)\\&\quad :\mathsf{comp}_2[\![x,y,f,z,g]\!]\to_{x\to_\star z}\mathsf{comp}_2[\![x,y,f,z,g]\!])[x,y,f,z,g]\end{aligned}$$

This term is "trying" to be the identity on $\mathsf{comp}_2[\![x, y, f, z, g]\!]$ (it is, in fact, provably equivalent to it in Catt), but is not actually a *syntactic* identity. Such terms are recognizable by the fact that they are coherences for which the type expression has a source and target which are equal. Henceforth we refer to them as *endomorphism coherences*, and our third rule ENDO sets them equal to the identities they duplicate, in the case at hand

$$\mathbb{1}_1[\![x, z, \mathsf{comp}_2[\![x, y, f, z, g]\!]]\!]$$

We point out that this third type of reduction has no apparent analog in the strictly unital theory of [5], which is otherwise closely related.

Properties of Definitional Equality We record here some basic facts about the definitional equality relation of $\mathsf{Catt}_{\mathsf{su}}$ introduced above.

Lemma 2. *Composition of substitutions is compatible with taking quotients:*

$$(\mu \circ \sigma) /\!/ \alpha = (\mu /\!/ \alpha) \circ \sigma$$

Lemma 3. *Let* $\sigma, \sigma' : \mathsf{Sub}$ *such that* $\sigma = \sigma'$. *For* $A : \mathsf{Type}$ *and* $t : \mathsf{Term}$, *we have:*

$$A[\![\sigma]\!] = A[\![\sigma']\!] \qquad\qquad t[\![\sigma]\!] = t[\![\sigma']\!]$$

Proof. Since substitution on types is given by structural induction (and the base case $A = \star$ is trivial), the first equation follows from the second.

Now, if t is a variable, then the result is clear by the definition of equality on substitutions, which is just equality of the comprising terms. On the other hand, $t \equiv \mathsf{coh}\,(\Gamma : A)[\tau]$, then we are reduced to showing that $\tau \circ \sigma = \tau \circ \sigma'$ and this is clear by expanding the definition and applying the induction hypothesis. □

Lemma 4. *Let* $\tau : \mathsf{Sub}$, $A, A' : \mathsf{Type}$ *and* $t, t' : \mathsf{Term}$. *If* $A = A'$ *and* $t = t'$, *then:*

$$A[\![\tau]\!] = A'[\![\tau]\!] \qquad\qquad t[\![\tau]\!] = t'[\![\tau]\!]$$

Proof. As in the previous Lemma, the first statement follows from the second by just inducting on the structure of the equality relation on types.

For the second, we argue by induction on the structure of the proof that $t = t'$. It suffices to check the generating cases, as the rest will follow by structural induction and the reflexivity, symmetry and transitivity of $=$.

Hence, suppose we have

$$t \equiv \mathsf{coh}\,(\Gamma : A)[\sigma] = \mathsf{coh}\,(\Gamma /\!/ \alpha : A[\![\pi_\alpha]\!])[\sigma /\!/ x] \equiv t'$$

Then we are reduced to showing that $(\sigma \circ \tau) /\!/ \alpha = (\sigma /\!/ \alpha) \circ \tau$ which is Lemma 2.

Next, if we have

$$t \equiv \mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\sigma] = d_{n+1}[\![\sigma]\!] \equiv t'$$

we argue as follows:

$$\begin{aligned} \mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\sigma][\![\tau]\!] &= \mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\sigma \circ \tau] \\ &= d_{n+1}[\![\sigma \circ \tau]\!] \\ &= d_{n+1}[\![\sigma]\!][\![\tau]\!] \end{aligned}$$

Finally, in the case that

$$t \equiv \mathsf{coh}\,(\Gamma : t \to_A t)[\sigma] = \mathbb{1}_{\mathsf{dim} A+1}[\![\{A[\![\sigma]\!]\}, t[\![\sigma]\!]]\!] \equiv t'$$

we obtain

$$\begin{aligned} \mathsf{coh}\,(\Gamma : t \to_A t)[\sigma][\![\tau]\!] &= \mathsf{coh}\,(\Gamma : t \to_A t)[\sigma \circ \tau] \\ &= \mathbb{1}_{\mathsf{dim} A+1}[\![\{A[\![\sigma \circ \tau]\!]\}, t[\![\sigma \circ \tau]\!]]\!] \\ &= \mathbb{1}_{\mathsf{dim} A+1}[\![\{A[\![\sigma]\!]\}, t[\![\sigma]\!]]\!][\![\tau]\!] \end{aligned}$$

where the last step follows from Lemma 1. □

Proposition 5. *If* $\Delta \vdash \mathsf{coh}\,(\Gamma : U)[\sigma] : A$ *in* $\mathsf{Catt_{su}}$*, then* $A = U[\![\sigma]\!]$*.*

Proof. We induction on the structure of the derivation of $\Delta \vdash \mathsf{coh}\,(\Gamma : U)[\sigma] : A$. Since the term in question is not a variable, there are only two cases. If the derivation is simply the introduction rule for coherences, then the result is immediate, as $U[\![\sigma]\!]$ is the assigned type.

Otherwise, the proof must be by the use of the conversion rule for terms. In this case, we obtain some Δ', A', Γ', U' and σ' together with the obvious equalities between them. By the induction hypothesis, we have that $A' = U'[\![\sigma']\!]$. And now, applying Lemmas 4 and 3 we have

$$A = A' = U'[\![\sigma']\!] = U'[\![\sigma]\!] = U[\![\sigma]\!]$$

which completes the proof. □

Inferred Types We note that Proposition 5 guarantees that if $\Delta \vdash \mathsf{coh}\,(\Gamma : u \to_T v)[\sigma] : A$ is valid, then $A = (u[\![\sigma]\!] \to_{T[\sigma]} v[\![\sigma]\!])$. So the type of a valid coherence term can be extracted from the syntax of the term itself up to definitional equality. We use this to define the *inferred type* of a coherence term, as follows:

$$\mathsf{ty}(\mathsf{coh}\,(\Gamma : U)[\sigma]) := U[\![\sigma]\!]$$

Furthermore, we define the *inferred source* and *inferred target* as follows:

$$\mathsf{src}(\mathsf{coh}\,(\Gamma : u \to_T v)[\sigma]) := u[\![\sigma]\!]$$
$$\mathsf{tgt}(\mathsf{coh}\,(\Gamma : u \to_T v)[\sigma]) := v[\![\sigma]\!]$$

We note that a term of the form $\mathsf{coh}\,(\Gamma : \star)[\ldots]$ is never valid, so $\mathsf{src}, \mathsf{tgt}$ are defined for every valid coherence term. Note that this notation is consistent with that introduced above for variables in a pasting context.

We also write $\mathsf{src}^k, \mathsf{tgt}^k$ for the iterated k-fold inferred source or target, and src_k, tgt_k for the k-dimensional inferred source or target; so for a coherence term t of dimension n, we have $\mathsf{src}_k(t) := \mathsf{src}^{n-k}(t)$, $\mathsf{tgt}_k(t) := \mathsf{tgt}^{n-k}(t)$.

2.3 Models of Catt and $\mathsf{Catt}_{\mathsf{su}}$

The type theories Catt and $\mathsf{Catt}_{\mathsf{su}}$ generate syntactic categories via a standard construction: objects are contexts, and morphisms are substitutions, up to definitional equality. Composition is given is via composition of substitutions. We abuse notation slightly and write Catt and $\mathsf{Catt}_{\mathsf{su}}$ for the corresponding categories. Furthermore, we write $\mathsf{Catt}^{\mathsf{pd}}$ and $\mathsf{Catt}^{\mathsf{pd}}_{\mathsf{su}}$ for the full subcategories consisting of only the pasting contexts.

As we have seen, this category contains a collection of objects D^k corresponding to the k-dimensional disc context. Taken together, these contexts, and their source and target substitutions constitute a *globular* object.

$$\cdots \underset{\tau}{\overset{\sigma}{\rightrightarrows}} \mathrm{D}^{k+1} \underset{\tau}{\overset{\sigma}{\rightrightarrows}} \mathrm{D}^{k} \underset{\tau}{\overset{\sigma}{\rightrightarrows}} \cdots \underset{\tau}{\overset{\sigma}{\rightrightarrows}} \mathrm{D}^{0}$$

Definition 6. A category $\mathcal{C}$ containing a globular object is said to admit *globular limits* if every diagram of the form

$$\mathrm{D}^{i_0} \xrightarrow{\tau^{l_0}} \mathrm{D}^{i_1} \xleftarrow{\sigma^{l_2}} \mathrm{D}^{i_2} \quad \cdots \quad \mathrm{D}^{i_{n-1}} \xleftarrow{\sigma^{l_n}} \mathrm{D}^{i_n}$$

admits a limit. Dually, a category $\mathcal{C}$ containing a co-globular object is said to admit *globular sums* if the category $\mathcal{C}^{\mathrm{op}}$ admits globular limits.

Theorem 7. *The categories* Catt *and* $\mathsf{Catt}_{\mathsf{su}}$ *admit globular limits.*

The proof that Catt admits globular limits is to some extent folklore, and has recently been written out [6]. The proof for $\mathsf{Catt}_{\mathsf{su}}$ would be precisely the same, as it only depends on the variable structure of pasting contexts, which is the same in both theories.

With this in hand, we are ready to give our notion of model.

Definition 8. An ∞-*category* is a presheaf on the category $\mathsf{Catt}^{\mathsf{pd}}$ which sends globular limits in $\mathsf{Catt}^{\mathsf{pd}}$ to globular sums in Set. A *strictly unital* ∞-*category* is a presheaf on $\mathsf{Catt}^{\mathsf{pd}}_{\mathsf{su}}$ which sends globular limits to globular sums.

In this notion of model of strictly unital ∞-category, algebraic operations are represented by morphisms of $\mathsf{Catt}^{\mathsf{pd}}_{\mathsf{su}}$, and strict unitality of those algebraic operations is then ensured in any model, since any definitionally-equal substitutions are identified in the category.

3 Reduction

Overview In this section we introduce a reduction relation $\rightsquigarrow$ on types, terms and substitutions, and show that its reflexive, transitive and symmetric closure agrees with definitional equality. We then define a subrelation called *standard reduction*, written $\boldsymbol{\rightsquigarrow}$, and show it is a partial function which terminates after finitely many steps, giving us a notion of normal form. Finally, we show that two terms have the same normal form just when they are definitionally equal, meaning that standard reduction gives an algorithm for deciding definitional equality.

Convention on Contexts For this section, and the remainder of the paper, we restrict our attention to contexts Γ which are pasting contexts; that is, for which $\Gamma \vdash_p$ is valid. When we speak of a valid type U, we mean a type for which there exists a pasting context Γ such that $\Gamma \vdash U$ is valid; and when we speak of a valid term t, we mean a term for which there is a pasting context Γ and type U such that $\Gamma \vdash t : U$ is valid. The reason for this restriction is that it simplifies our analysis—in particular, contexts do not reduce—and it is sufficient for our application to strictly unital ∞-categories. For valid coherence terms, we freely use Proposition 5 to infer information about its structure.

Reflexive, Transitive, Symmetric Closures Given a relation $\rightsquigarrow$, we write $\rightsquigarrow_{\mathsf{r}}$ for its reflexive closure, $\rightsquigarrow_{\mathsf{t}}$ for its transitive closure, and $\rightsquigarrow_{\mathsf{s}}$ for its symmetric closure. When we use multiple such subscripts, we mean this simultaneously; for instance, we write $\rightsquigarrow_{\mathsf{rts}}$ for its simultaneous reflexive, transitive and symmetric closure.

3.1 General reduction

We define a reduction relation on types, terms and substitutions, and show that the equivalence relation generated by these relations agrees with definitional equality.

We first define a simple syntactic property on terms, that of being an identity.

Definition 9. A term is an *identity* if it is of the form $\mathbb{1}_n[\![\sigma]\!]$ for some $n \in \mathbb{N}$; that is, when its head is an identity coherence.

We emphasise that as a syntactic property, this is not compatible with definitional equality. For example, if $t \equiv \mathbb{1}_n[\![\sigma]\!]$, then t is an identity; but if we merely have $t = \mathbb{1}_n[\![\sigma]\!]$, then t is not necessarily an identity.

We now give the reduction relation. The following definitions are given by simultaneous induction. A key point to observe is that determining whether a given type, term or substitution is a redex is a purely *syntactic* condition, which can be mechanically checked, and does not refer to definitional equality.

Definition 10 (Reduction of types)**.** The basic type $\star$ does not reduce. An arrow type $U \equiv (u \to_T v)$ reduces as follows:

(T1) if $u \rightsquigarrow u'$, then:

$$(u \to_T v) \overset{\text{T1}}{\rightsquigarrow} (u' \to_T v)$$

(T2) if $v \rightsquigarrow v'$, then:

$$(u \to_T v) \overset{\text{T2}}{\rightsquigarrow} (v \to_T v')$$

(T3) if $T \rightsquigarrow T'$, then:
$$(u \to_T v) \overset{\text{T3}}{\rightsquigarrow} (u \to_{T'} v)$$

Definition 11 (Reduction of substitutions)**.** A substitution $\sigma \equiv [s_1, \ldots, s_n]$ reduces as follows, given a reduction $s_i \rightsquigarrow s_i'$ of some argument:

$$[s_1, \ldots, s_i, \ldots, s_n] \overset{\text{S}}{\rightsquigarrow} [s_1, \ldots, s_i', \ldots, s_n]$$

Definition 12 (Reduction of terms)**.** Variable terms do not reduce. A coherence term $t \equiv \mathsf{coh}\,(\Gamma : T)[\sigma]$ reduces as follows:

(A) if $\sigma \rightsquigarrow \sigma'$, then:

$$\mathsf{coh}\,(\Gamma : T)[\sigma] \overset{\text{A}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : T)[\sigma']$$

(B) if t is not an identity, and $x \in \mathsf{var}(\Gamma)$ is a locally-maximal variable for which $x[\sigma]$ is an identity, then we define:

$$\mathsf{coh}\,(\Gamma : T)[\sigma] \overset{\text{B}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma /\!\!/ x : T[\![\pi_x]\!])[\sigma /\!\!/ x]$$

(C) if $T \rightsquigarrow T'$, then:

$$\mathsf{coh}\,(\Gamma : T)[\sigma] \overset{\text{C}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : T')[\sigma]$$

(D) the disc removal relation:

$$\mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, t] \overset{\text{D}}{\rightsquigarrow} t$$

(E) if t is not an identity, the endomorphism coherence removal relation (recall the notation $\{T\}$ from Section 2.2 of the unfolding of a type T):

$$\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] \overset{\text{E}}{\rightsquigarrow} \mathbb{1}_{\mathsf{dim} T+1}[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!]$$

If we can reduce $u \rightsquigarrow u'$ via some reduction stage (X) above, we say that u is a *general X-redex*, or just an *X-redex*, and we write $u \overset{\text{X}}{\rightsquigarrow} u'$. A given term can be a general X-redex for multiple stages (X). For example, if $u \rightsquigarrow u'$, then the term $\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma]$ is a C-redex in at least 2 ways, and also an E-redex, as follows:

$$\begin{aligned}
\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] &\overset{\text{C}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : u' \to_T u)[\sigma] \\
\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] &\overset{\text{C}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : u \to_T u')[\sigma] \\
\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] &\overset{\text{E}}{\rightsquigarrow} \mathbb{1}_{\mathsf{dim}_T+1}[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!]
\end{aligned}$$

A term could also have no reductions at all. So reduction is partially defined, and multivalued in general.

We now show that this equivalence relation generated by this reduction relation agrees with definitional equality constructed in Section 2. This is not trivial, because there are some key differences between how these relations are defined. Definitional equality allows pruning with respect to any argument $\alpha[\![\sigma]\!] = \mathbb{1}_n[\![\tau]\!]$ which is an identity *up to definitional equality*, but the reduction relation $\overset{\mathsf{B}}{\rightsquigarrow}$ only allows pruning with respect to an argument $\alpha[\![\sigma]\!] \equiv \mathbb{1}_n[\![\tau]\!]$ which is a *syntactic* identity. Definitional equality also allows endo-coherence removal for any $\mathsf{coh}\,(\Gamma : u \to_T v)[\sigma]$ with $u = v$, while E-reduction requires $u \equiv v$. Furthermore, the reduction relation explicitly proscribes identities as B- or E-redexes, but definitional equality has no such proscription.

Proposition 13. *On valid terms, types and substitutions, the equality relation $=$ agrees with $\rightsquigarrow_{\mathsf{rts}}$.*

Proof. Clearly $p \rightsquigarrow q$ implies $p = q$, since reduction is by definition a subrelation of equality, so we focus on the converse direction. Also, since the statement for types and substitutions immediately reduces to the statement for terms, we focus on terms here.

We say that a term p is *conservative* if for all terms q, we have that $p = q$ implies $p \rightsquigarrow_{\mathsf{rts}} q$. To prove the lemma, we must therefore show that all terms are conservative. Our proof operates by induction on subterms of p, and by case analysis on the equality $p = q$. Almost all such cases are immediate; here we explicitly handle the only nontrivial cases.

For the first nontrivial case, suppose $p = q$ is the following equality, obtained by pruning an identity term:

$$\begin{aligned} &\mathbb{1}_n[\![\ldots, v_2, v_1, v_1', \mathbb{1}_{n-1}[\![\ldots, u_3, u_2, u_1]\!]]\!] \\ &= (\mathsf{coh}\,\mathrm{D}^{n-1} : \mathbb{1}_{n-1} \to_{(d_{n-1} \to_{S^{n-2}} d_{n-1})} \mathbb{1}_{n-1})[\ldots, v_2, v_1] \end{aligned}$$

By induction on subterms of p we conclude that the terms u_i, v_1', v_i are conservative. By validity we must have $u_1 = v_1 = v_1'$ and $u_i\ v_i$, and hence we conclude $u_1 \rightsquigarrow_{\mathsf{rts}} v_1 \rightsquigarrow_{\mathsf{rts}} v_1'$ and $u_i \rightsquigarrow_{\mathsf{rts}} v_i$. We do not have $p \overset{\mathsf{B}}{\rightsquigarrow} q$, since p is an identity term, which are explicitly proscribed as B-redexes. However, q admits an E-reduction, and then a further series of A-reductions obtained by conservativity of the subterms of p, as follows:

$$\begin{aligned} q &\overset{\mathsf{E}}{\rightsquigarrow} \mathbb{1}_n[\![\ldots, v_3, v_2, v_1, v_1, \mathbb{1}_{n-1}[\![\ldots, v_3, v_2, v_1]\!]]\!] \\ &\overset{\mathsf{A}}{\rightsquigarrow}_{\mathsf{rts}} \mathbb{1}_n[\![\ldots, v_3, v_2, v_1, v_1', \mathbb{1}_{n-1}[\![\ldots, u_3, u_2, u_1]\!]]\!] \\ &\equiv p \end{aligned}$$

Hence $q \rightsquigarrow_{\mathsf{rts}} p$ as required.

For the second nontrivial case, suppose $p = q$ is the following equality, by the endo-coherence removal rule:

$$p \equiv \mathsf{coh}\,(\mathrm{D}^n : d_n \to_{S^{n-1}} d_n)[\ldots, u_2, u_1] = \mathbb{1}_n[\![d_n[\![\ldots, u_2, u_1]\!], \{S^{n-1}[\![\ldots, u_2, u_1]\!]\}]\!] \equiv q$$

We do not have $p \overset{\mathsf{E}}{\rightsquigarrow} q$, since p is an identity term, which is proscribed as an E-redex. However, we note that

$$q \equiv \mathbb{1}_n\left[\!\!\left[d_n[\![\ldots, u_2, u_1]\!], \{S^{n-1}[\![\ldots, u_2, u_1]\!]\}\right]\!\!\right] \equiv \mathbb{1}_n[\![\ldots, u_2, u_1]\!] \equiv p$$

So in fact $p \equiv q$, and hence $p \rightsquigarrow_{\mathsf{rts}} q$ as required.

For the third nontrivial case, we suppose $p = q$ is the following equality, with $\alpha = \mathbb{1}_n[\![\tau]\!]$ an argument in locally-maximal position:

$$p \equiv \mathsf{coh}\,(\Gamma : U)[\ldots, s, t, \alpha, \ldots] = \mathsf{coh}\,(\Gamma /\!\!/ x : U[\![\pi_x]\!])[\ldots, s, \ldots] \equiv q$$

By inductive hypothesis α is conservative, and since $\alpha = \mathbb{1}_n[\![\tau]\!]$ that yields the following:

$$p \overset{\text{A}}{\rightsquigarrow}_{\mathsf{rts}} \mathsf{coh}\,(\Gamma : U)[\langle \ldots, s, t, \mathbb{1}_n[\![\tau]\!], \ldots \rangle] \overset{\text{B}}{\rightsquigarrow} q$$

Hence $p \rightsquigarrow_{\mathsf{rts}} q$ as required.

For the fourth nontrivial case, we suppose $p = q$ is the following equality, where $u = u'$:

$$p \equiv \mathsf{coh}\,(\Gamma : u \to_T u')[\sigma] = \mathbb{1}_{\mathsf{dim}T+1}[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!] \equiv q$$

We do not necessarily have $u \equiv u'$, so cannot necessarily conclude $p \overset{\text{E}}{\rightsquigarrow} q$. But by induction on subterms we know u' is conservative, and thus from $u = u'$ we conclude $u' \rightsquigarrow_{\mathsf{rts}} u$. If $\mathsf{coh}\,(\Gamma : u \to_T u)$ is the head of an identity term, we proceed as for the second case above. Otherwise we argue as follows:

$$p \equiv \mathsf{coh}\,(\Gamma : u \to_T u')[\sigma] \overset{\text{C}}{\rightsquigarrow}_{\mathsf{rts}} \mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] \overset{\text{E}}{\rightsquigarrow} \mathbb{1}_{\mathsf{dim}T+1}[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!]$$

This completes the proof. □

Lemma 14. *Reduction of types, terms and substitutions preserves validity of judgements:*

- *if* $\Gamma \vdash A$ *is valid and* $A \rightsquigarrow A'$, *then* $\Gamma \vdash A'$ *is valid;*
- *if* $\Gamma \vdash t : A$ *is valid and* $t \rightsquigarrow t'$, *then* $\Gamma \vdash t' : A$ *is valid;*
- *if* $\Gamma \vdash t : A$ *is valid and* $A \rightsquigarrow A'$, *then* $\Gamma \vdash t : A'$ *is valid;*
- *if* $\Gamma \vdash \sigma : \Delta$ *is valid and* $\sigma \rightsquigarrow \sigma'$, *then* $\Gamma \vdash \sigma' : \Delta$ *is valid.*

Proof. Immediate since reduction is a subrelation of definitional equality, and validity can be transported across definitional equality (see Figure 13.) □

3.2 Standard reduction

We now define *standard reduction*, denoted with a bold arrow $\boldsymbol{\rightsquigarrow}$, a subrelation of general reduction $\rightsquigarrow$. Standard reduction has the property of being a reduction strategy, in the following sense.

Definition 15. A *reduction strategy* is a relation $\to$ on terms with the property that, if $a \to b$ and $a \to b'$, then $b \equiv b'$.

Put another way, a relation is a reduction strategy just when it is a partial function.

Standard reduction works in a similar way to reduction, but the reductions now have a preference order, so that higher-priority redexes, listed earlier in the following list, block lower-priority redexes listed later. Standard reduction is hence a reduction strategy by construction.

Definition 16 (Standard reduction of types)**.** The standard reduction of a type $U \equiv (u \to_T v)$ is given by the first matching reduction in the following list, if any:

(T0) if $T \rightsquigarrow \widetilde{T}$, then:
$$(u \to_T v) \overset{\text{T0}}{\rightsquigarrow} (u \to_{\widetilde{T}} v)$$

(T1) if $u \rightsquigarrow \tilde{u}$, then:
$$(u \to_T v) \overset{\text{T1}}{\rightsquigarrow} (\tilde{u} \to_T v)$$

(T2) if $v \rightsquigarrow \tilde{v}$, then:
$$(u \to_T v) \overset{\text{T2}}{\rightsquigarrow} (v \to_T \tilde{v})$$

Definition 17 (Standard reduction of substitutions)**.** Given a substitution $\sigma \equiv \langle s_1, \ldots, s_n \rangle$, then if $s_i \rightsquigarrow \tilde{s}_i$ is the leftmost argument with a standard reduction, we have the following:

$$[s_1, \ldots, s_i, \ldots, s_n] \overset{\text{S}}{\rightsquigarrow} [s_1, \ldots, \tilde{s}_i, \ldots, s_n]$$

Definition 18 (Standard reduction of terms)**.** A coherence term $t \equiv \mathsf{coh}\,(\Gamma : U)[\sigma]$ has a *standard reduction* given by the first relation in the following list which is defined, if any:

(A) if $\sigma \rightsquigarrow \tilde{\sigma}$, then:
$$\mathsf{coh}\,(\Gamma : U)[\sigma] \overset{\text{A}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : U)[\tilde{\sigma}]$$

(B) if t is not an identity, and $x \in \mathsf{var}(\Gamma)$ is the leftmost locally-maximal variable for which $x[\sigma]$ is an identity, then we define:
$$\mathsf{coh}\,(\Gamma : U)[\sigma] \overset{\text{B}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma /\!/ x : U[\![\pi_x]\!])[\sigma /\!/ x]$$

(C) if $T \rightsquigarrow \tilde{T}$, then:
$$\mathsf{coh}\,(\Gamma : T)[\sigma] \overset{\text{C}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : \tilde{T})[\sigma]$$

(D) the disc removal relation:
$$\mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, t] \overset{\text{D}}{\rightsquigarrow} t$$

(E) if t is not an identity, the endo-coherence removal relation:
$$\mathsf{coh}\,(\Gamma : u \to_A u)[\sigma] \overset{\text{E}}{\rightsquigarrow} \mathbb{1}_{\mathsf{dim}A+1}[\![\{A[\![\sigma]\!]\}, u[\![\sigma]\!]]\!]$$

If we can reduce $s \rightsquigarrow t$ via some reduction label (X) above, we say that u is a *standard X-redex*. It is an immediate consequence of the definition of standard reduction that it is a reduction strategy; that is, if a term, type or substitution has a standard reduction, it has *exactly one* standard reduction. This is quite unlike general reduction as defined as above. For example, suppose $u \rightsquigarrow \tilde{u}$, and consider the term $t \equiv \mathsf{coh}\,(\Gamma : u \to_A u)[\sigma]$. It is possible that t is a standard A-redex; failing that, it could be a standard B-redex; failing that, it will certainly be a standard C-redex. Although t is an E-redex (that is, there exists t' with $t \overset{\text{E}}{\rightsquigarrow} t'$), it is not a *standard* E-redex, since standard C-reductions are higher-priority than standard E-reductions.

Since standard reduction is unique when it exists, it is useful to introduce the following notation.

Definition 19. If s has a standard reduction, we write it as $\tilde{s}$, and hence $s \rightsquigarrow \tilde{s}$. We call $\tilde{s}$ the *standard reduct* of s.

We now record some results that show a term is reducible just when it is standard reducible.

Lemma 20 (Standard reductions are reductions)**.** *If $s \rightsquigarrow \tilde{s}$, then $s \rightsquigarrow \tilde{s}$.*

Proof. By definition, standard reduction is a subrelation of general reduction. □

Proposition 21 (Reducible terms are standard reducible)**.** *If $s \rightsquigarrow t$, then there exists a unique $\tilde{s}$ with $s \rightsquigarrow \tilde{s}$.*

Proof. Since standard reduction is a reduction strategy, uniqueness is clear. What we must establish is existence. The intuition is straightforward: in essence, we define standard reduction by giving a priority order to the redexes for general reduction, and allowing only the highest-priority redex. The result is then immediate, because if s has at least one reduction, then there must be a highest-priority such reduction.

We prove the result formally as follows, by simultaneous induction on the structure of terms, types, and substitutions. For the base cases, given by the type $\star$ or a variable term, there is no reduction, so the claim is vacuously true.

For a compound type $U \equiv (a \rightarrow_T b)$ is a type, then the statement follows immediately by induction on a, b or T.

For a substitution $\sigma = \langle s_1, \ldots, s_n \rangle$, suppose we have some reduction $\sigma \rightsquigarrow \tau$ arising from some choice of index i and some reduction $s_i \rightsquigarrow t$. Since σ has a reducible argument, it must have a leftmost reducible argument, which we can write as s_j, with $j \leq i$. By induction on subterms $s_j \rightsquigarrow \tilde{s}_j$, and $[\ldots, s_{j-1}, s_j, s_{j+1}, \ldots] \rightsquigarrow [\ldots, s_{j-1}, \tilde{s}_j, s_{j+1}, \ldots]$ is the required standard reduction.

For a coherence term $s \equiv \mathsf{coh}\,(\Gamma : T)[\sigma]$ with a reduction $s \rightsquigarrow t$, we argue by case analysis as follows.

- If u is an A-redex, there must exist some σ' such that $\sigma \rightsquigarrow \sigma'$. By induction $\sigma \rightsquigarrow \tilde{\sigma}$, and hence $\mathsf{coh}\,(\Gamma : T)[\sigma] \overset{\mathrm{A}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : T)[\tilde{\sigma}]$.
- If u is not an A-redex, but u is a B-redex, then there must be some leftmost locally-maximal argument of Γ with respect to which it is a standard B-redex.
- If u is not an A- or B-redex, but it is a C-redex, then there must exist some T' such that $T \rightsquigarrow T'$. Hence by induction $T \rightsquigarrow \tilde{T}$, and so $u \overset{\mathrm{C}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : \tilde{T})[\sigma]$.
- If u is not an A-, B- or C-redex, but it is a D-redex, then the D-reduction will be standard.
- If u is not an A-, B-, C- or D-redex, but it is an E-redex, then the E-reduction will be standard.
- If u is not an A-, B-, C-, D- or E-redex, then u cannot be reduced, contradicting the hypothesis of the theorem.

This completes the argument. □

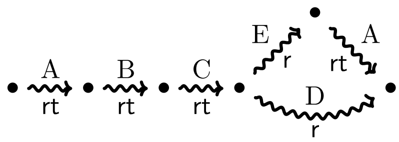


Figure 14: Standard reduction pathways to normal form

3.3 Termination of standard reduction

Standard reduction gives us a notion of normal form, as follows.

Definition 22. We define a term, type or substitution to be in *normal form* when it has no reduction, either by standard or general reduction (by Lemmas 20 and 21, these conditions are equivalent.)

In this section we show that standard reduction terminates after finite time on every term, type and substitution. This means that for every term t we can obtain a term $N(t)$ in normal form, by repeatedly applying standard reduction until a normal form is reached.

To work towards our termination result, we consider what happens when we start with a given term and repeatedly perform standard reduction. What we find is a specific pattern of standard reductions, illustrated in Figure 14. We first perform some family of standard A-reductions. If these terminate, they will be followed by some family of standard B-reductions. If these terminate, they will be followed by some family of standard C-reductions. If these terminate, and they are not yet sufficient to yield a normal form, they will be followed either by a single standard D-reduction, giving a term in normal form; or by a single standard E-reduction and a family of standard A-reductions, which if they terminate will yield a normal form. We prove this claim as follows.

Proposition 23. *The reflexive transitive closure of the standard reduction relation is obtained as the following composite:*

$$\rightsquigarrow_{\mathsf{rt}} \quad = \quad \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}} \overset{\mathrm{B}}{\rightsquigarrow}_{\mathsf{rt}} \overset{\mathrm{C}}{\rightsquigarrow}_{\mathsf{rt}} \left(\overset{\mathrm{D}}{\rightsquigarrow}_{\mathsf{r}} \cup \overset{\mathrm{E}}{\rightsquigarrow}_{\mathsf{r}} \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}} \right)$$

Proof. We establish the claim by showing that the following composite reduction pairs and triples are impossible (that is, they are empty as relations).

- $p \overset{\mathrm{B}}{\rightsquigarrow} q \overset{\mathrm{A}}{\rightsquigarrow} r$. For p to be a standard B–redex, it is required that it is not an A-redex, and so the arguments of p must be in normal form. But the arguments of q are a subset of the arguments of p, contradicting the claim that q is a standard A-redex.
- $p \overset{\mathrm{C}}{\rightsquigarrow} q \overset{\mathrm{A}}{\rightsquigarrow} r$. For p to be a standard C-redex, it is required that it is not an A-redex, and so the arguments of p must be in normal form. But C-reductions do not change the arguments of a term, so the arguments of q are the same as the arguments p, contradicting the claim that q is a standard A-redex.
- $p \overset{\mathrm{C}}{\rightsquigarrow} q \overset{\mathrm{B}}{\rightsquigarrow} r$. For p to be a standard C-redex, it is required that it is not a B-redex, a condition that depends on the context and arguments of the term. But C-reductions do not change the context or arguments, so it is impossible that q is a B-redex.

- $p \overset{\mathrm{D}}{\rightsquigarrow} q \rightsquigarrow r$. For a term to be a standard D-redex, it is required that it is not an A-redex, and so the arguments of p must not be reducible. By the action of D-reduction, the term q is one of the arguments of p, contradicting the claim that q is reducible.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{B}}{\rightsquigarrow} r$. The standard E-reduct q is an identity term, but identity terms are never B-redexes, by definition.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{C}}{\rightsquigarrow} r$. The standard E-reduct q is an identity term, but identity terms are never C-redexes, since the cell part is in normal form.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{D}}{\rightsquigarrow} r$. The standard E-reduct q is an identity term, but identity terms are never D-redexes, as the head has the wrong syntactic form.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{E}}{\rightsquigarrow} r$. The standard E-reduct q is an identity term, but identity terms are never E-redexes, by definition.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}} r \overset{\mathrm{B}}{\rightsquigarrow} s$. The standard E-reduct q is an identity term. A-reductions do not change the head, and so r will also be an identity term. But identity terms are never B-redexes, giving a contradiction.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}} r \overset{\mathrm{C}}{\rightsquigarrow} s$. The standard E-reduct q is an identity term. A-reductions do not change the head, and so r will also be an identity term. But identity terms are never C-redexes, giving a contradiction.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}} r \overset{\mathrm{D}}{\rightsquigarrow} s$. The standard E-reduct q is an identity term. A-reductions do not change the head, and so r will also be an identity term. But identity terms are never D-redexes, giving a contradiction.

- $p \overset{\mathrm{E}}{\rightsquigarrow} q \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}} r \overset{\mathrm{E}}{\rightsquigarrow} s$. The standard E-reduct q is an identity term. A-reductions do not change the head, and so r will also be an identity term. But identity terms are never E-redexes, by definition.

The result is then established as follows, by imagining a standard reduction sequence for some given coherence term. Here we refer to composite relations by concatenation; so for example, $t \overset{\mathrm{B}\,\mathrm{A}}{\rightsquigarrow\rightsquigarrow} t'$ just when there exists some t'' with $t \overset{\mathrm{B}}{\rightsquigarrow} t'' \overset{\mathrm{A}}{\rightsquigarrow} t'$.

- Standard A-reductions have the highest priority, so these will be performed first.

- If the above step terminates, standard B-reductions have the second-highest priority, so we now perform these. Since $\overset{\mathrm{B}\,\mathrm{A}}{\rightsquigarrow\rightsquigarrow}$ is empty, this will not trigger any additional standard A-reductions.

- If the above step terminates, standard C-reductions have the third-highest priority, so we now perform these. Since $\overset{\mathrm{C}\,\mathrm{A}}{\rightsquigarrow\rightsquigarrow}$ and $\overset{\mathrm{C}\,\mathrm{B}}{\rightsquigarrow\rightsquigarrow}$ are both empty, these standard C-reductions will not trigger further standard A- or B-reductions.

- If the above step terminates, standard D-reductions have the fourth-highest priority. If we can perform a D-reduction, the result will be in normal form, since $\overset{\mathrm{D}}{\rightsquigarrow}\rightsquigarrow$ is empty.

- If we cannot perform a D-reduction, we consider applying a standard E-reduction, as the standard reduction with fifth-highest priority. If the standard E-reduction cannot be applied, then the term is in normal form, as the standard E-reduction is the last reduction in the list.

- If the standard E-reduction was successfully applied, it cannot be followed by a standard B-, C-, D-, or E- reduction, since $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{B}}{\rightsquigarrow}$, $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{C}}{\rightsquigarrow}$, $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{D}}{\rightsquigarrow}$ and $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{E}}{\rightsquigarrow}$ are all shown above to be empty. The only remaining possibility is for the standard E-reduction to be followed by some sequence of standard A-reductions. These A-reductions cannot themselves be followed by a standard B-, C-, D- or E-reduction, since we show above that $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}}\overset{\mathrm{B}}{\rightsquigarrow}$, $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}}\overset{\mathrm{C}}{\rightsquigarrow}$, $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}}\overset{\mathrm{D}}{\rightsquigarrow}$ and $\overset{\mathrm{E}}{\rightsquigarrow}\overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rt}}\overset{\mathrm{E}}{\rightsquigarrow}$ are all empty.

This completes the proof. ☐

We next show that standard reduction is terminating.

Proposition 24. *Standard reduction is terminating on valid types, terms and substitutions.*

Proof. Standard reduction on types and substitutions is given in terms of standard reduction of a finite family of terms, so we need only check that standard reduction of terms has no infinite sequences. The variable case is trivial, so we consider reduction of some coherence term $t \equiv \mathsf{coh}\,(\Gamma : T)[\sigma]$.

We proceed by simultaneous induction, on the dimension of t, and on subterms of t. Since a subterm of a valid term can never have a higher dimension, this is well-defined. The dimension is not defined for variables, but since variables are in normal form, this does not affect the argument.

Thanks to Proposition 23, we know that the standard reduction of a given term can be separated into finitely many distinct phases of standard A-, B-, C-, D- and E-reduction, with each phase involving standard reductions of a single fixed type. So we need only show that each standard reduction phase will terminate.

- *Standard A-reduction.* By induction on subterms, since a substitution is of finite length, a given term will have only finitely many standard A-reductions.

- *Standard B-reduction.* Since the context has a finite number of variables, a given term will have only finitely many standard B-reductions.

- *Standard C-reduction.* Since the cell part of a valid term is formed from terms of strictly lower dimension, it follows by induction on dimension that a term will have only finitely many standard C-reductions.

Standard D- and E-reductions are single-step operations, so no termination argument is necessary for those. ☐

Having now established that standard reduction has no infinite sequences, it is clear every term has a unique normal form with respect to it.

Definition 25. For any term t, its *normal form* $N(t)$ is the unique term with $t \rightsquigarrow_{\mathsf{rt}} N(t)$ such that there is no term u with $N(t) \rightsquigarrow u$.

Definition 26. For any term t, its *distance* is the length, possibly zero, of the unique standard reduction sequence $t \rightsquigarrow t_1 \rightsquigarrow \cdots \rightsquigarrow t_n \rightsquigarrow t'$ to its normal form.

3.4 Technical results on reduction

Here we collect further results on reduction, mostly of a technical nature, which will be used in the next subsection for our main proof.

Lemma 27. *Term substitution is compatible with substitution reduction:*

$$\sigma \rightsquigarrow \sigma' \quad \Rightarrow \quad u[\![\sigma]\!] \rightsquigarrow_{\mathsf{rt}} u[\![\sigma']\!]$$

Proof. We induct on the structure of u. If u is a variable, then either $u[\![\sigma]\!] \equiv \sigma'[\![\tau]\!]$ or $u[\![\sigma]\!] \rightsquigarrow u[\![\sigma']\!]$; in either case, we have $u[\![\sigma]\!] \rightsquigarrow_{\mathsf{r}} u[\![\sigma']\!]$ as required.

Otherwise, we have $u \equiv \mathsf{coh}\,(\Gamma : U)[\rho]$, and we argue as follows:

$$\begin{aligned}
u[\sigma] &\equiv \mathsf{coh}\,(\Gamma : U)[\rho \circ \sigma] \\
&\equiv \mathsf{coh}\,(\Gamma : U)[\rho_1[\![\sigma]\!], \rho_2[\![\sigma]\!], \ldots, \rho_n[\![\sigma]\!]] \\
&\overset{\mathsf{A}}{\rightsquigarrow}_{\mathsf{rt}} \mathsf{coh}\,(\Gamma : U)[\rho_1[\![\sigma']\!], \rho_2[\![\sigma]\!], \ldots, \rho_n[\![\sigma]\!]] \\
&\overset{\mathsf{A}}{\rightsquigarrow}_{\mathsf{rt}} \mathsf{coh}\,(\Gamma : U)[\rho_1[\![\sigma']\!], \rho_2[\![\sigma']\!], \ldots, \rho_n[\![\sigma]\!]] \\
&\overset{\mathsf{A}}{\rightsquigarrow}_{\mathsf{rt}} \cdots \\
&\overset{\mathsf{A}}{\rightsquigarrow}_{\mathsf{rt}} \mathsf{coh}\,(\Gamma : U)[\rho_1[\![\sigma']\!], \rho_2[\![\sigma']\!], \ldots, \rho_n[\![\sigma']\!]] \\
&\equiv \mathsf{coh}\,(\Gamma : U)[\rho \circ \sigma'] \\
&\equiv u[\![\sigma']\!]
\end{aligned}$$

Hence $u[\![\sigma]\!] \rightsquigarrow_{\mathsf{rt}} u[\![\sigma']\!]$ as required. □

Proof. If u is a variable, no reduction is possible, in contradiction with hypothesis. We therefore assume $u \equiv \mathsf{coh}\,(\Gamma : T)[\mu]$ is a coherence term, writing $\mu = [\mu_1, \ldots, \mu_n]$, and proceed by case analysis on the structure of the reduction $u \rightsquigarrow u'$, and by induction on subterms of u.

If $u \overset{\mathsf{A}}{\rightsquigarrow} u'$ via some $m_i \rightsquigarrow m_i'$, then we argue as follows:

$$\begin{aligned}
u[\![\sigma]\!] &\equiv \mathsf{coh}\,(\Gamma : U)[\mu \circ \sigma] \\
&\equiv \mathsf{coh}\,(\Gamma : U)[\mu_1[\![\sigma]\!], \ldots, \mu_i[\![\sigma]\!], \ldots, \mu_k[\![\sigma]\!]] \\
&\overset{\mathsf{A}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : U)[\mu_1[\![\sigma]\!], \ldots, \mu_i'[\![\sigma]\!], \ldots, \mu_k[\![\sigma]\!]] \\
&\equiv u[\![\sigma']\!]
\end{aligned}$$

Alternatively, we suppose $u \overset{\mathsf{B}}{\rightsquigarrow} u'$ is the B-reduction $\mathsf{coh}\,(\Gamma : U)[\mu] \overset{\mathsf{B}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma /\!/ x : U[\![\pi_x]\!])[\mu /\!/ x]$, eliminating some locally-maximal variable x_i of Γ for which $x_i[\![\mu]\!] \equiv \mu_i$ is an identity. Then $\mu_i[\![\sigma]\!]$ is also an identity, and hence using Lemma 2 we have:

$$\begin{aligned}
u[\![\sigma]\!] &\equiv \mathsf{coh}\,(\Gamma : U)[\mu \circ \sigma] \\
&\equiv \mathsf{coh}\,(\Gamma : U)[\mu \circ \sigma] \\
&\overset{\mathsf{B}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma /\!/ x : U[\![\pi_x]\!])[(\mu \circ \sigma) /\!/ x] \\
&\equiv \mathsf{coh}\,(\Gamma /\!/ x : U[\![\pi_x]\!])[(\mu /\!/ x) \circ \sigma] \\
&\equiv u'[\![\sigma]\!]
\end{aligned}$$

If $u \overset{\mathsf{C}}{\rightsquigarrow} u'$ via some $T \overset{\mathsf{C}}{\rightsquigarrow} T'$, then we argue as follows:

$$u[\![\sigma]\!] \equiv \mathsf{coh}\,(\Gamma : T)[\mu \circ \sigma] \overset{\mathsf{C}}{\rightsquigarrow} \mathsf{coh}\,(\Gamma : T')[\mu \circ \sigma] \equiv u'[\![\sigma]\!]$$

If $u \overset{\mathbb{D}}{\rightsquigarrow} u'$ as $\mathsf{coh}\,(\mathrm{D}^n : S^{n-1})[\ldots, u'] \overset{\mathbb{D}}{\rightsquigarrow} u'$, then we argue as follows:

$$u[\![\sigma]\!] \equiv \mathsf{coh}\,(\mathrm{D}^n : S^{n-1})[\ldots, u'[\![\sigma]\!]] \overset{\mathbb{D}}{\rightsquigarrow} u'[\![\sigma]\!]$$

Finally, if $u \overset{\mathbb{E}}{\rightsquigarrow} u'$ as $\mathsf{coh}\,(\Gamma : u \to_U u)[\mu] \overset{\mathbb{E}}{\rightsquigarrow} \mathbb{1}_{\mathsf{dim}U+1}\big[\!\big[\{U[\![\mu]\!]\}, u[\![\mu]\!]\big]\!\big]$, then we argue as follows, using Lemma 1:

$$\begin{aligned} u[\![\sigma]\!] &\equiv \mathsf{coh}\,(\Gamma : u \to_U u)[\mu \circ \sigma] \\ &\overset{\mathbb{E}}{\rightsquigarrow} \mathbb{1}_{\mathsf{dim}U+1}\big[\!\big[\{U[\![\mu \circ \sigma]\!]\}, u[\![\mu \circ \sigma]\!]\big]\!\big] \\ &\equiv \mathbb{1}_{\mathsf{dim}U+1}[\![\{U[\![\mu]\!]\}, u[\![\mu]\!]]\!][\![\sigma]\!] \\ &\equiv u'[\![\sigma]\!] \end{aligned}$$

This completes the proof. □

Lemma 28. *Term substitution is compatible with term reduction:*

$$u \rightsquigarrow u' \quad \Rightarrow \quad u[\![\sigma]\!] \rightsquigarrow u'[\![\sigma]\!]$$

Lemma 29 (Identities reduce to identities)**.** *If u is an identity, and $u \rightsquigarrow u'$, then u' is an identity.*

Proof. We recognize an identity term by looking at the head. We prove the result by case analysis on the reduction $u \rightsquigarrow u'$. If $u \overset{\mathbb{A}}{\rightsquigarrow} u'$ the result is immediate, since A-reductions do not change the head. If $u \overset{\mathbb{B}}{\rightsquigarrow} u'$ we have a contradiction, since identity terms cannot be B-redexes by definition. If $u \overset{\mathbb{C}}{\rightsquigarrow} u'$ we again have a contradiction, since a C-reduction acts on the head of the term, but the head of an identity term is in normal form. If $u \overset{\mathbb{D}}{\rightsquigarrow} u'$ we again have a contradiction, since identity terms have the wrong form to be D-redexes. If $u \overset{\mathbb{E}}{\rightsquigarrow} u'$ we again have a contradiction, since identity terms cannot be E-redexes by definition. □

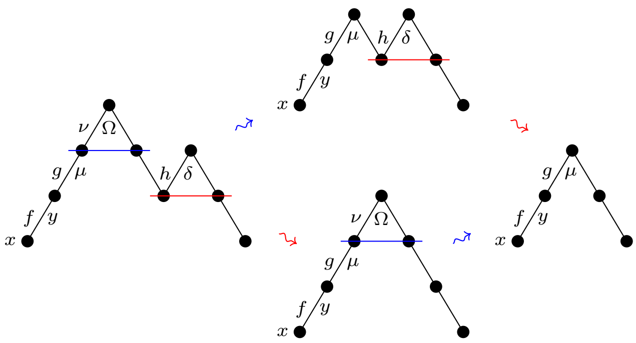
Lemma 30. *If $\sigma \rightsquigarrow \tilde{\sigma}$, then $\sigma /\!\!/ x \rightsquigarrow_{\mathsf{r}} \tilde{\sigma} /\!\!/ x$.*

Proof. This is immediate, since $\sigma /\!\!/ x$ is a sublist of σ. If the first reducible argument of σ is not in the sublist $\sigma /\!\!/ x$, then $\sigma /\!\!/ x \equiv \tilde{\sigma} /\!\!/ x$. Otherwise, it will still be the first reducible argument of the sublist, and $\sigma /\!\!/ x \rightsquigarrow \tilde{\sigma} /\!\!/ x$. □

Lemma 31. *Given a context Γ and distinct locally-maximal variables x, y, and a substitution $\Delta \vdash \sigma : \Gamma$ with $x[\![\sigma]\!], y[\![\sigma]\!]$ both identities, the following contexts and substitutions are identical:*

$$(\Gamma /\!\!/ x) /\!\!/ y \equiv (\Gamma /\!\!/ y) /\!\!/ x \qquad \pi_x \circ \pi_y \equiv \pi_y \circ \pi_x \qquad (\sigma /\!\!/ x) /\!\!/ y \equiv (\sigma /\!\!/ y) /\!\!/ x$$

Proof. The first statement is clear from example, in this case showing $(\Gamma /\!\!/ \Omega) /\!\!/ \delta \equiv (\Gamma /\!\!/ \delta) /\!\!/ \Omega$, since the excision operations are independent:



The other claims follows similarly. □

Lemma 32. *For a valid substitution $\Delta \vdash \sigma : \mathrm{D}^n$, we have for all $i \leq n$:*

$$d_i[\![\sigma]\!] = \mathsf{src}^{n-i}(d_n[\![\sigma]\!]) \qquad\qquad d'_i[\![\sigma]\!] = \mathsf{tgt}^{n-i}(d_n[\![\sigma]\!])$$

Proof. Extending the substitution construction rule along definitional equality we must have $\mathsf{src}(d_n[\![\sigma]\!]) = \mathsf{src}(d_n)[\![\sigma]\!]$, and similarly for tgt. The result follows. □

Lemma 33. *For a valid substitution σ that sends a locally-maximal variable x to an identity, we have $\pi_x \circ (\sigma /\!/ x) = \sigma$.*

Proof. Suppose $x[\sigma] = \mathbb{1}_n[\![\tau_0, \tau'_0, \ldots, \tau_{n-1}, \tau'_{n-1}, \tau_n]\!]$, an identity, and suppose $\mathsf{src}(x) = v$ and $\mathsf{tgt}(x) = w$ as variables. Then by the substitution construction rule transported along definitional equality, $\mathsf{src}(\mathbb{1}_n[\![\tau]\!]) = \mathsf{tgt}(\mathbb{1}_n[\![\tau]\!]) = \tau_n$. So $\sigma = [\ldots, \tau_n, \tau_n, \mathbb{1}_n[\![\tau]\!], \ldots]$, and furthermore $\sigma /\!/ x = [\ldots, \tau_n, \ldots]$, and $\pi_x \equiv [\ldots, v, v, \mathbb{1}_n[\![\mathsf{src}^n(v), \mathsf{tgt}^n(v), \ldots, \mathsf{src}(v), \mathsf{tgt}(v), v]\!], \ldots]$. Composing, we obtain:

$$\begin{aligned}
\pi_x \circ (\sigma /\!/ x) &\equiv [\ldots, v[\![\sigma /\!/ x]\!], v[\![\sigma /\!/ x]\!], \\
&\qquad \mathbb{1}_n[\![\mathsf{src}^n(v), \mathsf{tgt}^n(v), \ldots, \mathsf{src}(v), \mathsf{tgt}(v), v]\!][\![\sigma /\!/ x]\!], \ldots] \\
&= [\ldots, \tau_n, \tau_n, \mathbb{1}_n[\![\mathsf{src}^n(\tau_n), \mathsf{tgt}^n(\tau_n) \ldots, \mathsf{src}(\tau_n), \mathsf{tgt}(\tau_n), \tau_n]\!], \ldots] \\
&= [\ldots, \tau_n, \tau_n, \mathbb{1}_n[\![\tau_0, \tau'_0, \ldots, \tau_{n-1}, \tau'_{n-1}, \tau_n]\!], \ldots] \\
&= \sigma
\end{aligned}$$

The penultimate step here uses Lemma 32. This completes the proof. □

3.5 Standard reduction generates definitional equality

We show that the symmetric, transitive and reflexive closure of standard reduction generates definitional equality. Since we have already shown that standard reduction is a terminating reduction strategy, this gives an algorithm to determine whether two given terms are definitionally equal, by computing their standard normal forms and checking if they are syntactically equal.

Lemma 34. *If $a \rightsquigarrow_{\mathsf{rts}} b$, there exists c with $a \rightsquigarrow_{\mathsf{rt}} c$ and $b \rightsquigarrow_{\mathsf{rt}} c$.*

Proof. This is immediate, since $\rightsquigarrow$ is a reduction strategy. □

Proposition 35. *We have the following, for any valid term s:*

(i) If $s \rightsquigarrow t$, we can find terms a, b which admit a reduction $a \rightsquigarrow_{\mathsf{r}} b$ and standard reductions $s \rightsquigarrow_{\mathsf{t}} a$, $t \rightsquigarrow_{\mathsf{rts}} b$, illustrated as follows:

$$\begin{array}{ccc}
s & \rightsquigarrow & t \\
{\scriptstyle \mathsf{t}}\downarrow & & \downarrow{\scriptstyle \mathsf{rts}} \\
a & \underset{\mathsf{r}}{\rightsquigarrow} & b
\end{array}$$

(ii) If $s = t$, then $s \rightsquigarrow_{\mathsf{rts}} t$.

Proof. Since by Proposition 24 every term reaches normal form after a finite number of standard reductions, and since by Lemma 13 reduction generates equality, it is clear that statement (i) implies statement (ii). We therefore focus here on the proof of (i).

We can neglect the case of s being a variable, since variables do not reduce. It follows that s is a coherence term, and since the dimension of a valid coherence term is well-defined, we will make use of that throughout. The proof of (i) is by simultaneous induction on the dimension of s, and on subterms of s. Since no subterm of s has a greater dimension than s itself, this is consistent.

One possibility, which arises several times in the case analysis below, is that the reduction $s \rightsquigarrow t$ is itself standard (that is, $t \equiv \tilde{s}$.) We can handle this case once-and-for-all as follows:

$$\begin{array}{ccc} s & \rightsquigarrow & t \\ \Big\downarrow & & \Big\| \\ \tilde{s} & = & \tilde{s} \end{array} \tag{1}$$

We refer to this argument below where it is needed.

We now begin the main proof of property (i), by case analysis on the reduction $s \rightsquigarrow t$.

First case $s \overset{A}{\rightsquigarrow} t$. We suppose $s \equiv \mathsf{coh}\,(\Gamma : T)[\sigma]$, and that $s \overset{A}{\rightsquigarrow} t$ by reducing an argument of σ via $s_i \rightsquigarrow s_i'$. Then s must be a standard A-redex, because if some argument is not in normal form, there must exist a leftmost argument not in normal form; so we have $s \overset{A}{\rightsquigarrow} \tilde{s}$.

Suppose $s \overset{A}{\rightsquigarrow} t$ and $s \overset{A}{\rightsquigarrow} \tilde{s}$ act by reducing the same argument of s_i of σ, via $s_i \rightsquigarrow s_i'$ and $s_i \rightsquigarrow \tilde{s}_i$ respectively. If $s_i' \equiv \tilde{s}_i$ then we are done by the argument above expression (1) above. Otherwise the result holds by induction on the subterm s_i, as follows:

$$\begin{array}{ccc} \mathsf{coh}\,(\Gamma : T)[\ldots, s_i, \ldots] & \overset{A}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : T)[\ldots, s_i', \ldots] \\ {\scriptstyle A}\Big\downarrow & & {\scriptstyle \mathsf{rts}}\Big\downarrow{\scriptstyle A} \\ \mathsf{coh}\,(\Gamma : T)[\ldots, \tilde{s}_i, \ldots] & \underset{A}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : T)[\ldots, q, \ldots] \end{array}$$

Alternatively, suppose they act by reducing different arguments of σ. Then the redexes are independent, and we argue as follows:

$$\begin{array}{ccc} \mathsf{coh}\,(\Gamma : T)[\ldots, s_i, \ldots, s_j, \ldots] & \overset{A}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : T)[\ldots, s_i, \ldots, s_j', \ldots] \\ {\scriptstyle A}\Big\downarrow & & \Big\downarrow{\scriptstyle A} \\ \mathsf{coh}\,(\Gamma : T)[\ldots, \tilde{s}_i, \ldots, s_j, \ldots] & \underset{A}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : T)[\ldots, \tilde{s}_i, \ldots, s_j', \ldots] \end{array}$$

Second case $s \overset{B}{\rightsquigarrow} t$. For this case, we suppose $s \overset{B}{\rightsquigarrow} t$ as follows:

$$s \equiv \mathsf{coh}\,(\Gamma : T)[\sigma] \overset{B}{\rightsquigarrow} \mathsf{coh}\,(\Gamma /\!\!/ x : T[\![\pi_x]\!])[\sigma /\!\!/ x] \equiv t$$

We proceed by case analysis on the standard reduction $s \rightsquigarrow \tilde{s}$.

- *Standard A-reduction.* In this case, we argue as follows:

$$\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : T)[\sigma] & \overset{\mathrm{B}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma /\!\!/ x : T[\![\pi_x]\!])[\sigma /\!\!/ x] \\
\mathrm{A}\,\big\downarrow & & \big\downarrow\,\mathrm{A} \\
\mathsf{coh}\,(\Gamma : T)[\tilde{\sigma}] & \underset{\mathrm{B}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma /\!\!/ x : T[\![\pi_x]\!])[\tilde{\sigma} /\!\!/ x]
\end{array}$$

Since the upper B-reduction is valid, we know that $x[\![\sigma]\!]$ is an identity; then by Lemma 29 we also have that $x[\![\tilde{\sigma}]\!]$ is an identity, and so the lower B-reduction is valid. Validity of the standard A-reduction on the right of the square follows from Lemma 30.

- *Standard B-reduction.* We suppose s is a standard B-redex with respect to some locally-maximal variable y. If $x \equiv y$, then the reductions are the same, and this case is handled by the general argument above expression (1). Otherwise, we argue as follows, using Lemma 31:

$$\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : T)[\sigma] & \overset{\mathrm{B}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma /\!\!/ x : T[\![\pi_x]\!])[\sigma /\!\!/ x] \\
 & & \big\downarrow\,\mathrm{B} \\
\mathrm{B}\,\big\downarrow & & \mathsf{coh}\,\big((\Gamma /\!\!/ x) /\!\!/ y : T[\![\pi_x]\!][\![\pi_y]\!]\big)[(\sigma /\!\!/ x) /\!\!/ y] \\
 & & \Vert \\
\mathsf{coh}\,(\Gamma /\!\!/ y : T[\![\pi_y]\!])[\sigma /\!\!/ y] & \underset{\mathrm{B}}{\rightsquigarrow} & \mathsf{coh}\,\big((\Gamma /\!\!/ y) /\!\!/ x : T[\![\pi_y]\!][\![\pi_x]\!]\big)[(\sigma /\!\!/ y) /\!\!/ x]
\end{array}$$

Third case $s \overset{\mathrm{C}}{\rightsquigarrow} t$. Suppose $s \equiv \mathsf{coh}\,(\Gamma : T)[\sigma]$, and $s \overset{\mathrm{C}}{\rightsquigarrow} t$ acts via some type reduction $T \rightsquigarrow U$. Then we argue by case analysis on the reduction $s \rightsquigarrow \tilde{s}$.

- *Standard A-reduction.* In this case we have the following:

$$\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : T)[\sigma] & \overset{\mathrm{C}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : U)[\sigma] \\
\mathrm{A}\,\big\downarrow & & \big\downarrow\,\mathrm{A} \\
\mathsf{coh}\,(\Gamma : T)[\tilde{\sigma}] & \underset{\mathrm{C}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : U)[\tilde{\sigma}]
\end{array}$$

- *Standard B-reduction.* If s is a standard B-redex, we have the following, employing Lemma 28:

$$\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : T)[\sigma] & \overset{\mathrm{C}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : U)[\sigma] \\
\mathrm{B}\,\big\downarrow & & \big\downarrow\,\mathrm{B} \\
\mathsf{coh}\,(\Gamma /\!\!/ x : T[\![\pi_x]\!])[\sigma /\!\!/ x] & \underset{\mathrm{C}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma /\!\!/ x : U[\![\pi_x]\!])[\sigma /\!\!/ x]
\end{array}$$

- *Standard C-reduction.* Suppose s is a standard C-redex via a standard type reduction $T \rightsquigarrow \tilde{T}$. If $\tilde{T} \equiv U$, we are done by the argument above expression (1).

Otherwise, since $T \rightsquigarrow U$, by induction on subterms we know that $U \rightsquigarrow_{\mathsf{rts}} T$, and we argue as follows:

$$
\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : T)[\sigma] & \overset{\mathrm{C}}{\rightsquigarrow} & \mathsf{coh}\,(\Gamma : U)[\sigma] \\
 & & \mathsf{rts}\downarrow \mathrm{C} \\
\mathrm{C}\downarrow \mathsf{t} & & \mathsf{coh}\,(\Gamma : T)[\tau] \\
 & & \downarrow \mathrm{C} \\
\mathsf{coh}\,(\Gamma : \tilde{T})[\tau] & = & \mathsf{coh}\,(\Gamma : \tilde{T})[\tau]
\end{array}
$$

Fourth case $s \overset{\mathrm{D}}{\rightsquigarrow} t$*.* We suppose $s \overset{\mathrm{D}}{\rightsquigarrow} t$ as follows:

$$s \equiv \mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, t] \overset{\mathrm{D}}{\rightsquigarrow} t$$

Then we consider the standard reduction for s.

- *Standard A-reduction.* Supposing $t \rightsquigarrow \tilde{t}$ is the leftmost reducible argument of s, then we have the following:

$$
\begin{array}{ccc}
\mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, t] & \overset{\mathrm{D}}{\rightsquigarrow} & t \\
\mathrm{A}\downarrow & & \downarrow \\
\mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, \tilde{t}] & \underset{\mathrm{D}}{\rightsquigarrow} & \tilde{t}
\end{array}
$$

Otherwise, let p be the leftmost reducible argument of s. Then we argue as follows:

$$
\begin{array}{ccc}
\mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, p, \ldots, t] & \overset{\mathrm{D}}{\rightsquigarrow} & p \\
\mathrm{A}\downarrow & & \| \\
\mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, \tilde{p}, \ldots, t] & \underset{\mathrm{D}}{\rightsquigarrow} & p
\end{array}
$$

- *Standard B-reduction.* In this case we must have $t \equiv \mathbb{1}_n[\![\ldots, q_2, q_1]\!]$, and hence $s \equiv \mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, p_2, p_1, p_1', \mathbb{1}_n[\![\ldots, q_2, q_1]\!]]$. Since s is valid we deduce $q_1 = p_1 = p_1'$ and $q_i = p_i$. It follows by induction on subterms that $q_i \rightsquigarrow_{\mathsf{rts}} p_i$. We put this together as follows:

$$
\begin{array}{ccc}
\mathsf{coh}\,(\mathrm{D}^{n+1} : S^n)[\ldots, p_2, p_1, p_1', \mathbb{1}_n[\![\ldots, q_2, q_1]\!]] & \overset{\mathrm{D}}{\rightsquigarrow} & \mathbb{1}_n[\![\ldots, q_2, q_1]\!] \\
\mathrm{B}\downarrow & & \mathsf{rts}\downarrow \mathrm{A} \\
\mathsf{coh}\,(\mathrm{D}^n : d_n \to_{S^{n-1}} d_n)[\ldots, p_2, p_1] & = & \mathbb{1}_n[\![\ldots, p_2, p_1]\!]
\end{array}
$$

- *Standard C-reduction.* The term s cannot be C-redex, since the type S^n is in normal form, being constructed entirely from variables.
- *Standard D-reduction.* In this case, the result follows from the argument around around expression (1) above.

Fifth case $s \overset{\mathrm{E}}{\rightsquigarrow} t$*.* We suppose $s \overset{\mathrm{E}}{\rightsquigarrow} t$ as follows, for $n = \mathsf{dim} T + 1$:

$$s \equiv \mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] \rightsquigarrow \mathbb{1}_n[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!] \equiv t$$

We now consider the structure of the standard reduction $s \rightsquigarrow \tilde{s}$.

- *Standard A-reduction.* If $\sigma \rightsquigarrow \tilde{\sigma}$ then by Lemma 27 we have $u[\![\sigma]\!] \rightsquigarrow_{\mathsf{rt}} u[\![\tilde{\sigma}]\!]$ and $T[\![\sigma]\!] \rightsquigarrow_{\mathsf{rt}} T[\![\tilde{\sigma}]\!]$. We conclude by induction on dimension that $\mathbb{1}_n[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!] \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rts}} \mathbb{1}_n[\![\{T[\![\tilde{\sigma}]\!]\}, u[\![\tilde{\sigma}]\!]]\!]$. Altogether, we have the following as required:

$$\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] & \overset{\mathrm{E}}{\rightsquigarrow} & \mathbb{1}_n[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!] \\
\mathrm{A}\,\wr & & \mathsf{rts}\,\wr\,\mathrm{A} \\
\mathsf{coh}\,(\Gamma : u \to_T u)[\tilde{\sigma}] & \underset{\mathrm{E}}{\rightsquigarrow} & \mathbb{1}_n[\![\{T[\![\tilde{\sigma}]\!]\}, u[\![\tilde{\sigma}]\!]]\!]
\end{array}$$

- *Standard B-reduction.* In this case we have the following:

$$\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] & \overset{\mathrm{E}}{\rightsquigarrow} & \mathbb{1}_n\left[\!\!\left[\{T[\![\sigma]\!]\}, u[\![\sigma]\!]\right]\!\!\right] \\
\mathrm{B}\,\wr & & \mathsf{rts}\,\wr\,\mathrm{A} \\
\mathsf{coh}\,(\Gamma /\!\!/ x : u[\![\pi_x]\!] \to_{T[\![\pi_x]\!]} u[\![\pi_x]\!])[\sigma /\!\!/ x] & \underset{\mathrm{E}}{\rightsquigarrow} & \mathbb{1}_n\left[\!\!\left[\{T[\![\pi_x]\!][\![\sigma /\!\!/ x]\!]\}, u[\![\pi_x]\!][\![\sigma /\!\!/ x]\!]\right]\!\!\right]
\end{array}$$

We obtain the right-hand standard A-reduction as follows. From Lemma 33 we know $\sigma = \pi_x \circ (\sigma /\!\!/ x)$, and hence $u[\![\sigma]\!] = u[\![\pi_x]\!][\![\sigma /\!\!/ x]\!]$ and $T[\![\sigma]\!] = T[\![\pi_x]\!][\![\sigma /\!\!/ x]\!]$. By induction we conclude $u[\![\sigma]\!] \rightsquigarrow_{\mathsf{rts}} u[\![\pi_x]\!][\![\sigma /\!\!/ x]\!]$ and $T[\![\sigma]\!] \rightsquigarrow_{\mathsf{rts}} T[\![\pi_x]\!][\![\sigma /\!\!/ x]\!]$.

- *Standard C-reduction.* Since $u[\![\sigma]\!] = u'[\![\sigma]\!]$ and $T[\![\sigma]\!] = T'[\![\sigma]\!]$, it follows by induction that $u[\![\sigma]\!] \rightsquigarrow_{\mathsf{rts}} u'[\![\sigma]\!]$ and $T[\![\sigma]\!] \rightsquigarrow_{\mathsf{rts}} T'[\![\sigma]\!]$, and hence that $\mathbb{1}_n[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!] \overset{\mathrm{A}}{\rightsquigarrow}_{\mathsf{rts}} \mathbb{1}_n[\![\{T'[\![\sigma]\!]\}, u'[\![\sigma]\!]]\!]$. We then have the following:

$$\begin{array}{ccc}
\mathsf{coh}\,(\Gamma : u \to_T u)[\sigma] & \overset{\mathrm{E}}{\rightsquigarrow} & \mathbb{1}_n[\![\{T[\![\sigma]\!]\}, u[\![\sigma]\!]]\!] \\
\mathrm{C}\,\wr\,\mathsf{t} & & \mathsf{rts}\,\wr\,\mathrm{A} \\
\mathsf{coh}\,(\Gamma : N(u) \to_{N(T)} N(u))[\sigma] & \underset{\mathrm{E}}{\rightsquigarrow} & \mathbb{1}_n[\![\{N(T)[\![\sigma]\!]\}, N(u)[\![\sigma]\!]]\!]
\end{array}$$

- *Standard D-reduction.* The term s cannot be a D-redex, as the term s has an incompatible structure.
- *Standard E-reduction.* In this case we are done, with the result following from the argument around around expression (1).

This completes the argument. □

Corollary 36. *We have $s = t$ if and only if $N(s) \equiv N(t)$.*

3.6 Examples

Now that we can decide definitional equality, we can investigate some examples. These will illustrate the theory, and in particular allow us to show it is non-trivial (that is, there is not a single definitional equality class of terms.) To begin with, consider the context describing the three composable 1-morphisms:

$$\Gamma = (x : \star)(y : \star)(f : x \to_\star y)(z : \star)(g : y \to_\star z)(w : \star)(h : z \to_\star w)$$

Then the following three terms are all in normal form, but distinct, demonstrating the non-triviality of the theory:

- $\mathsf{comp}_3[\![x, y, f, z, g, w, h]\!]$, which represents $f \cdot g \cdot h$, the unbiased 3-fold composite;
- $\mathsf{comp}_2[\![x, z, \mathsf{comp}_2[\![x, y, f, z, g]\!], w, h]\!]$, which represents $(f \cdot g) \cdot h$, a biased binary composite;
- $\mathsf{comp}_2[\![x, y, f, w, \mathsf{comp}_2[\![y, z, g, w, h]\!]]\!]$, which represents $f \cdot (g \cdot h)$, another biased binary composite.

To see the reductions of the theory working, let us first normalize the following term, showing that $f \cdot \mathbb{1}_y$ normalizes to f:

$$\mathsf{comp}_2[\![x, y, f, y, \mathbb{1}_0[\![y]\!]]\!] \overset{\mathrm{B}}{\rightsquigarrow} \mathsf{comp}_1[\![x, y, f]\!] \overset{\mathrm{D}}{\rightsquigarrow} f$$

This is one of the examples we considered in Section 2; here we see that our standard reduction relation can algorithmically extract the normal form. For our second example, we show that $\lambda_f : f \cdot \mathbb{1}_y \Rightarrow f$ normalizes to id_f:

$$\begin{aligned}
\mathsf{unit\text{-}r} &\equiv \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y) : \mathsf{comp}_2[\![x, y, f, y, \mathbb{1}_0[\![y]\!]]\!] \to_{x \to_\star y} f)[x, y, f] \\
&\overset{\mathrm{C}}{\rightsquigarrow} \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y) : \mathsf{comp}_1[\![x, y, f]\!] \to_{x \to_\star y} f)[x, y, f] \\
&\overset{\mathrm{C}}{\rightsquigarrow} \mathsf{coh}\,((x : \star)(y : \star)(f : x \to_\star y) : f \to_{x \to_\star y} f)[x, y, f] \\
&\equiv \mathbb{1}_1[\![x, y, f]\!]
\end{aligned}$$

These examples illustrate the way that standard reduction lets us find the strictly unital normal form for any ∞-categorical operation, expressed as a $\mathsf{Catt}_{\mathsf{su}}$ term.

4 Disc trivialization

In the following section, we prove a structure theorem that says in a disc context D^n, up to definitional equality, every term is either a variable, or an iterated identity on a variable. So if we restrict to terms that use *all* variables of D^n (that is, the terms which do not factor through a smaller context), then there is exactly one definitional equivalence class of term in each dimension $k \geq n$, the normal form of which is the locally-maximal variable d_n or an iterated identity on this. In this sense, $\mathsf{Catt}_{\mathsf{su}}$ trivializes disc contexts.

This provides an interesting point of comparison with work of Batanin, Cisinski and Weber [5], discussed further in Section 1.2, where strictly unital weak ∞-categories are defined as algebras over a certain operad, defined to be the universal one which trivializes the operations over discs, and which satisfies a unit compatibility property. By showing our theory also trivializes operations over discs, we claim that in this respect, our theory is at least as strict.

Preparatory Lemmas Before we can prove the main theorem of this section, we need to establish some technical results about pasting contexts.

We say that a substitution $\Delta \vdash \sigma : \Gamma$ is a *variable-to-variable substitution* if for all $x \in \mathsf{var}(\Gamma)$, we have that $x[\![\sigma]\!]$ is again a variable.

Lemma 37. *Let $\mathrm{D}^n \vdash \sigma : \Gamma$ be a valid substitution which is in normal form, and which sends locally-maximal variables of Γ to variables of D^n. Then σ is a variable-to-variable substitution.*

Proof. Let v be a variable of Γ. Then there is some locally-maximal variable w of Γ such that $v = \mathsf{src}^k(w)$ for some $k \in \mathbb{N}$. It follows from the formation rules for substitution that $\mathsf{src}^k(w[\![\sigma]\!]) = \mathsf{src}^k(w)[\![\sigma]\!] \equiv v[\![\sigma]\!]$. Since we are given that $w[\![\sigma]\!]$ is a variable, it follows that $v[\![\sigma]\!]$ is a variable up to definitional equivalence. But since σ is normalized, $v[\![\sigma]\!]$ must be precisely a variable. $\square$

For every $k < n$, there are two variable-to-variable substitutions $\mathrm{D}^n \vdash \partial^{\pm}_{k,n} : \mathrm{D}^k$, which map the k-disc context into the appropriate source or target context of D^n. We also have $\mathrm{D}^n \vdash \mathrm{id}_{\mathrm{D}^n} : \mathrm{D}^n$, the identity substitution. We call these *subdisc inclusions.* We now show that every valid variable-to-variable substitution $\mathrm{D}^n \vdash \sigma : \Gamma$ is of this form.

Lemma 38. *Let* $\mathrm{D}^n \vdash \sigma : \Gamma$ *be a valid variable-to-variable substitution. Then we have* $\Gamma \equiv \mathrm{D}^k$ *for some* $k \leq n$*, and* σ *is a subdisc inclusion.*

Proof. The variables of a pasting context Γ form a globular set $\mathsf{g}(\Gamma)$ in an obvious way, and the substitution well-typedness condition means that a variable-to-variable substitution $\mathrm{D}^n \vdash \sigma : \Gamma$ induces a function of globular sets $\sigma : \mathsf{g}(\Gamma) \to \mathsf{g}(\mathrm{D}^n)$. Suppose for a contradiction Γ is *not* a disc: then it must contain some sub–Dyck word $(\Uparrow\ (\Downarrow\ (\Uparrow\ (\cdots)\, y\, f) \cdots) z\, g \cdots)$, and we have $\mathsf{tgt}(f) = \mathsf{src}(g)$. Then also $\mathsf{tgt}(\sigma(f)) = \mathsf{src}(\sigma(g))$; but now we have a contradiction, since the globular set of a disc does not have any pair of elements related in this way.

We conclude that $\Gamma \equiv \mathrm{D}^k$ for some $k \leq n$. It remains to show that σ is a subdisc inclusion. For this, suppose $k = n$. Then since σ preserves variable dimension, we must have $d_n[\![\sigma]\!] = d_n$, and this extends uniquely to the other variables, since σ is a function of globular sets, and we conclude $\sigma \equiv \mathrm{id}_{\mathrm{D}^n}$. Otherwise, suppose $k < n$. Then we can choose $d_k[\![\sigma]\!] = d_k$ or $d_k[\![\sigma]\!] = d'_k$, and once again, both extend uniquely, yielding $\sigma \equiv \partial^-_{k,n}$ and $\sigma \equiv \partial^+_{k,n}$ respectively. $\square$

Given a valid term t in some context Γ, its *canonical identity* is $\mathbb{1}(t) := \mathbb{1}_{\mathsf{dim}(t)}[\![\{\mathsf{ty}(t)\}, t]\!]$. Canonical identities can be distinguished from ordinary identities $\mathbb{1}_n[\![\sigma]\!]$ because we do not need to give the dimension subscript, as it can be inferred from the term and the supplied context; because we use round brackets; and because we supply a term as an argument, rather than a substitution. A term is an *iterated canonical identity* if it is of the form $\mathbb{1}^k(t)$, by applying this construction k times for $k > 0$. We now show that if a term is definitionally equal to an ordinary identity $\mathbb{1}_n[\![\sigma]\!]$, it is definitionally equal to a canonical identity.

Lemma 39. *If* t *is a valid term of* Γ *with* $t = \mathbb{1}_n[\![\ldots, p]\!]$*, then* $t = \mathbb{1}(p)$*.*

Proof. We define $\sigma := [\ldots, p_2, p'_2, p_1, p'_1, p]$. Because $\Gamma \vdash \sigma : \mathrm{D}^n$ is a valid substitution, it must satisfy the substitution typing conditions up to definitional equality, so we conclude for each $0 < k \leq n$ the following:

$$p_k \equiv d_{n-k}[\![\sigma]\!] \equiv \mathsf{src}^k(d_n)[\![\sigma]\!] = \mathsf{src}^k(d_n[\![\sigma]\!]) \equiv \mathsf{src}^k(p)$$
$$p'_k \equiv d'_{n-k}[\![\sigma]\!] \equiv \mathsf{tgt}^k(d_n)[\![\sigma]\!] = \mathsf{tgt}^k(d_n[\![\sigma]\!]) \equiv \mathsf{tgt}^k(p)$$

We now reason as follows:

$$t \equiv \mathbb{1}_n[\![\ldots, p_2, p'_2, p_1, p'_1, p]\!]$$

$$= \mathbb{1}_n[\![\ldots, \mathsf{src}^2(p), \mathsf{tgt}^2(p), \mathsf{src}(p), \mathsf{tgt}(p), p]\!]$$
$$\equiv \mathbb{1}(p)$$

This completes the proof. □

Structure Theorem We are now prepared to prove our structure theorem characterizing definitional equality for valid terms over discs, which says that any term of D^n in normal form must be syntactically equal to $\mathbb{1}^k(d)$ for some $k \in \mathbb{N}$ and some variable d of D^n.

Theorem 40 (Disc trivialization)**.** *Suppose t is valid in D^n. Then t is definitionally equal to a variable, or to the iterated canonical identity on a variable.*

Proof. If t is a variable, we are done. Otherwise, t is a coherence term, and we have that $t \equiv \mathsf{coh}\,(\Gamma : U)[\sigma]$. By Corollary 36, we may assume without loss of generality that t is in normal form.

If t is an identity, then by Lemma 39 we know $t = \mathbb{1}(u)$. By induction on dimension, u is therefore either a variable or a iterated identity on a variable, and we are done.

It remains to consider the case that $t \equiv \mathsf{coh}\,(\Gamma : U)[\sigma]$ is not an identity. We will see that we are now guaranteed to obtain a contradiction. Since t is in normal form, we know that t is not an A-, B-, C-, D- or E-redex, and we use these facts freely below.

First, note that $\mathrm{D}^n \vdash \sigma : \Gamma$ maps locally maximal variables of Γ to non-identity terms of D^n (or else t would be a B-redex), and these terms are in normal form (or else t would be an A-redex). Hence, by induction on subterms, we may assume that σ maps locally-maximal variables to variables. By Lemma 37, it follows that σ is a variable-to-variable substitution, and then from Lemma 38 we conclude that Γ is a disc context D^k with $k \leq n$, and σ is a subdisc inclusion. We therefore conclude that $t \equiv \mathsf{coh}\,(\mathrm{D}^k : u \to_T v)[\sigma]$.

Suppose $\mathsf{dim}(t) = k$. Then u, v must each use all the variables of the respective boundary context, so by induction on subterms, the only possibility is $u = d_{k-1}$ and $v = d'_{k-1}$. Since t is not a C-redex, we conclude that u is in normal form (hence $u \equiv d_{k-1}$), v is in normal form (hence $v \equiv d'_{k-1}$), and T is in normal form (hence $T \equiv S^{k-2}$), and so $t \equiv \mathsf{coh}\,(\mathrm{D}^k : d_{k-1} \to_{S^{k-2}} d'_{k-1})[\sigma]$. But then t would be a D-redex, which is a contradiction.

So we must have $\mathsf{dim}(t) > k$. Then u, v must each use all the variables of D^k, so by induction on subterms, the only possibility is $u = v = \mathbb{1}^{n-k}(d_k)$. But this would mean that t is an E-redex, again giving a contradiction. □

5 Rehydration

Overview The results of Section 3 tell us that normal forms decide definitional equality, in the following sense:

$$u = v \quad \Rightarrow \quad N(u) \equiv N(v)$$

Since normalization contracts definitional equivalence classes, and since Catt has a trivial notion of definitional equality, we might imagine this is sufficient to produce our encoding function ρ into Catt terms. However, a normalized valid $\mathsf{Catt_{su}}$ term t is *not* necessarily a valid Catt term. To see why, we note that the property above does not extend to boundaries of terms:

$$\mathsf{tgt}(u) = \mathsf{src}(v) \quad \not\Rightarrow \quad \mathsf{tgt}(N(u)) \equiv \mathsf{src}(N(v))$$

As a result, if a pair of normalized terms of $\mathsf{Catt_{su}}$ are composable, implicit coercion may *still* be required by the type checker to verify that the composite itself type checks.

The solution is to introduce a new notion called *rehydrated normal form*, written $R(N(t))$. Since it is a function of the normal form, it of course also decides definitional equality:

$$u = v \quad \Rightarrow \quad R(N(u)) \equiv R(N(v))$$

However, it also recursively composes the boundaries of all subterms with coherences, which has the effect of putting those boundaries *themselves* in rehydrated normal form. These coherences can be regarded as "explicit coercions", giving a direct structural witness to the change in type. As a result, we gain the following desirable feature:

$$\mathsf{tgt}(u) = \mathsf{src}(v) \quad \Rightarrow \quad \mathsf{tgt}(R(N(u)) \equiv \mathsf{src}(R(N(v))$$

Since it also acts recursively on the term structure, we are therefore guaranteed that any composable subterms of $R(N(t))$ will be composable on-the-nose. As a result, for any valid $\mathsf{Catt_{su}}$ term t, the term $R(N(t))$ will be a valid Catt term, all necessary coercions having been explicitly inserted.

Finally, since the coercions we insert all normalize to the identity, we are guaranteed that $R(N(t)) = t$ in $\mathsf{Catt_{su}}$, meaning that the definitional equivalence classes are preserved. If we define $\rho(t) := R(N(t))$, we therefore obtain the encoding function discussed in the introduction, providing a solution to the Identification Paradox for this "toy model" theory.

Rehydrated Normal Form We introduce the following operations on a term t simultaneously by mutual recursion:

- the *rehydration* $R(t)$, which rehydrates all subterms, and then pads the resulting term;
- the *padding* $P(t)$, which composes a term at its boundaries to ensure all of its sources and targets are in rehydrated normal form;
- the *normalizer* $\phi(t)$, a coherence term which provides an explicit equivalence between t and its rehydrated normal form $R(N(t))$.

We first give the definition of rehydration, in terms of the padding operation.

Definition 41. For a valid term t of $\mathsf{Catt_{su}}$, its *rehydration* $R(t)$ is defined as follows:

- $R(x) := x$

- $R(\mathsf{coh}\,(\Gamma : U)[\sigma]) := P(\mathsf{coh}\,(\Gamma : R(U))[R(\sigma)])$

On valid types and substitution, rehydration is defined by applying term rehydration to all subterms and subtypes.

We note that $\mathsf{supp}(t) = \mathsf{supp}(R(t))$ for all terms t; this follows since padding is the only non-recursive part of the rehydration operation, which composes the term with normalizers, which cannot change the support, as we discuss below underneath Definition 43.

We next give the definition of padding, in terms of the normalizers of its inferred sources and targets. Since these sources and targets have strictly smaller dimension than the term itself, the mutual recursion between rehydration, padding and normalizers is well-founded. We use the notation $\mathsf{comp}_{\mathsf{n},\mathsf{k}}$ for the coherence introduced on page 11 and encoding the "unbiased" composite of a n-dimensional disc D with two $(k+1)$-dimensional discs S and T glued along the k-dimensional source and target of D, respectively. This coherence is valid in a context with exactly three locally maximal cells corresponding to D, S and T. It will be convenient for the presentation in this section, therefore, to write this coherence by providing just these three arguments, i.e. $\mathsf{comp}_{\mathsf{n},\mathsf{k}}(a, b, c)$, regarding the non-locally maximal arguments as implicit.

Definition 42. For a valid $\mathsf{Catt}_{\mathsf{su}}$ term $\Gamma \vdash t : A$, its *padding* $P(t)$ is obtained by composing at the boundaries to put them into rehydrated normal form. We define $P(t) := P_{\mathsf{dim}A+1}(t)$, and then:

- $P_0(t) := t$
- $P_{k+1}(t) := \mathsf{comp}_{\mathsf{dim(A)+1,k}}(\phi(\mathsf{src}_k(t)), P_k(t), \phi^{-1}(\mathsf{tgt}_k(t)))$

The constructors P_k each "fix up" the corresponding inferred source and targets of their arguments, so $P_k(t)$ is guaranteed to have $\mathsf{src}_j(P_k(t))$ and $\mathsf{tgt}_j(P_k(t))$ in rehydrated normal form for all $j < k$.

Finally we give the definition of the normalizer, in terms of rehydration.

Definition 43. For a term t of $\mathsf{Catt}_{\mathsf{su}}$, valid over some pasting context Γ with $\mathsf{supp}(t) = \mathsf{FV}(\Gamma)$, its *normalizer* $\phi(t)$ is the following coherence:

$$\phi(t) := \mathsf{coh}\left(\Gamma : R(N(t)) \to_{\mathsf{ty}(u)} t\right)[\mathrm{id}_\Gamma]$$

Its *inverse* $\phi^{-1}(t)$ is defined with the reversed type, as follows:

$$\phi^{-1}(t) := \mathsf{coh}\left(\Gamma : t \to_{\mathsf{ty}(u)} R(N(t))\right)[\mathrm{id}_\Gamma]$$

By the formation rules for coherences, we note that $\mathsf{supp}(\mathsf{src}(\phi(t))) = \mathsf{supp}(\mathsf{tgt}(\phi(t))) = \mathsf{FV}(\Gamma)$. For this reason, composing a term with a normalizer does not change the support of that term.

Properties of Rehydrated Normal Forms We now verify that terms in rehydrated normal form have the required properties.

Lemma 44. *The inferred sources and targets of a term in rehydrated normal form are also in rehydrated normal form.*

Proof. We demonstrate this as follows:

$$\begin{aligned}\mathsf{src}(R(N(t)) &\equiv \mathsf{src}\Big(\mathsf{comp}_{\mathsf{dim}(A)+1,k}\big(\phi(\mathsf{src}_k(t)), P_k(t), \phi^{-1}(\mathsf{tgt}_k(t))\big)\Big)\\ &\equiv \mathsf{src}(\phi(\mathsf{src}_k(t)))\\ &\equiv \mathsf{src}(\mathsf{coh}\,(\Gamma : R(N(\mathsf{src}_k(t))) \to \mathsf{src}_k(t))[\mathrm{id}_\Gamma])\\ &\equiv R(N(\mathsf{src}_k(t)))\end{aligned}$$

This completes the proof. □

Corollary 45. *If* $\mathsf{tgt}^k(R(N(u))) = \mathsf{src}^j(R(N(v)))$ *are valid definitionally-equal terms, then they are syntactically equal.*

Proof. This is immediate, since we showed above that these boundaries are in rehydrated normal form, and if two terms in rehydrated normal form are definitionally equal, they must be syntactically equal. □

We now show that a rehydrated normal form $R(N(t))$ a valid Catt term. A valid $\mathsf{Catt}_{\mathsf{su}}$ term t will be valid over a number of contexts Γ; for simplicity, we suppose for the purposes of this theorem that we have chosen such a context with $\mathsf{supp}(t) = \mathsf{FV}(\Gamma)$, that is, for which Γ has no extraneous variables. Such a choice can always be made, although we do not prove that formally here.

Theorem 46. *If t is a valid term of* $\mathsf{Catt}_{\mathsf{su}}$*, over some pasting context Γ with* $\mathsf{supp}(t) = \mathsf{FV}(\Gamma)$*, then $R(N(t))$ is a valid term of* Catt.

Proof. We argue by induction on subterms, and by dimension on terms, which is consistent in this case, since a valid term cannot have a subterm of greater dimension. For variables, the claim is immediate, and we therefore focus on coherence terms $t \equiv \mathsf{coh}\,(\Gamma : U)[\sigma]$. Supposing such a term is in normal form, its rehydration is as follows:

$$R(t) \equiv P(\mathsf{coh}\,(\Gamma : R(U))[R(\sigma)])$$

The subterms of $R(U)$ and $R(\sigma)$ will be individually in rehydrated normal form, and hence by induction valid in Catt. Since R does not change the support of the term, $R(U)$ will satisfy the free variable condition for coh term formation, as required.

We next argue that the subterm $t' := \mathsf{coh}\,(\Gamma : R(U))[R(\sigma)]$ is valid in Catt. Validity of a $\mathsf{Catt}_{\mathsf{su}}$ term is generated inductively by typing inferences, as described in Section 2. In every case, the hypotheses for these inferences are that certain typing assertions for subterms are valid. From Proposition 5 we know that $\Gamma \vdash t : A$ just when $A = \mathsf{ty}(t)$; and so these typing hypotheses can be reduced to hypotheses that certain inferred boundary terms are definitionally equal to others. But for the term t', all subterms are in rehydrated normal form, and we show above in Corollary 45 that inferred boundaries of such terms are definitionally equal just when they are syntactically equal. Validity of the term therefore holds in Catt, which differs from $\mathsf{Catt}_{\mathsf{su}}$ only in having a restricted notion of definitional equality. Since t' is valid in Catt, its inferred sources and targets will also be valid in Catt.

We now argue that $P(t')$ is valid in Catt. We prove this by induction on the parameter k of the padding construction:

$$P_{k+1}(t) \equiv \mathsf{comp}_{\mathsf{dim(A)+1,k}}\big(\phi(\mathsf{src}_k(t)), P_k(t), \phi^{-1}(\mathsf{tgt}_k(t))\big)$$

The composition operations $\mathsf{comp}_{\mathsf{d,k}}$ are certainly valid Catt terms, as elementary syntactic constructions. By induction, the subterm $P_k(t)$ is valid in Catt. The normalizers in this expression are being computed for terms $\mathsf{src}_k(t)$,$\mathsf{tgt}_k(t)$. While these terms are not themselves in rehydrated normal form, their inferred types $\mathsf{ty}(\mathsf{src}_k(t))$, $\mathsf{ty}(\mathsf{tgt}_k(t))$ *will* be in rehydrated normal form, thanks to the remark following Definition 42.

To complete the proof, we must therefore show that if u is some valid term with $\mathsf{ty}(u)$ in rehydrated normal form, then $\phi(u)$ and $\phi^{-1}(u)$ are valid in Catt. We consider the definition of $\phi(u)$:

$$\phi(u) := \mathsf{coh}\left(\Gamma : R(N(u)) \to_{\mathsf{ty}(u)} u\right)[\mathrm{id}_\Gamma]$$

Since $\mathsf{dim}(u) < \mathsf{dim}(t)$, we know by induction on dimension that $R(N(u))$ is valid in Catt. We know u arises as an inferred source or target of t', and hence is valid in Catt. We also know $\mathsf{ty}(u)$ is already in rehydrated normal form, and hence $\mathsf{ty}(u) \equiv \mathsf{ty}(R(N(u)))$. So $\phi(u)$ is valid in Catt, as is $\phi^{-1}(u)$. □

Finally, we show that the rehydrated normal form is in the same definitional equivalence class of $\mathsf{Catt}_{\mathsf{su}}$ of the original term.

Proposition 47. *If t is a valid term of $\mathsf{Catt}_{\mathsf{su}}$ over some pasting context, then:*

$$R(N(t)) = t$$

Proof. For a variable, this is immediate, and we therefore focus on coherence terms. We may assume that the result holds on all terms u of strictly smaller dimension than t. This allows us to show that such terms have normalizers which are definitionally equal to identities:

$$\begin{aligned}
\phi(u) &\equiv \mathsf{coh}\,(\Gamma : u \to_{\mathsf{ty}(u)} R(N(u)))[\mathrm{id}_\Gamma] \\
&\overset{\text{C}}{\rightsquigarrow}_{\mathsf{rt}} \mathsf{coh}\,(\Gamma : N(u) \to_{N(\mathsf{ty}(u))} N(u))[\mathrm{id}_\Gamma] \\
&\overset{\text{E}}{\rightsquigarrow} \mathbb{1}_{\mathsf{dim}(t)}[\![\{N(\mathsf{ty}(u))\}, N(t)]\!]
\end{aligned}$$

Note that we use the equation $N(R(N(u))) \equiv N(u)$ in the first step, this being a consequence of the conclusion of the theorem applied to u, which is of smaller dimension than t.

Now, since paddings are constructed from the normalizers of terms of strictly smaller dimension (namely the sources and targets of t), we may use the previous result to show that paddings are definitionally equal to the term being padded. This we prove by induction on the parameter k in the definition of the padding composite P_k:

$$\begin{aligned}
P_{k+1}(t) &\equiv \mathsf{comp}_{\mathsf{n,k}}(\phi(\mathsf{src}_k(t)), P_k(t), \phi(\mathsf{tgt}_k(t))) \\
&\overset{\text{A}}{\rightsquigarrow}_{\mathsf{t}} \mathsf{comp}_{\mathsf{n,k}}(\mathbb{1}_k[\![\tau]\!], t, \mathbb{1}_k[\![\tau]\!]) \\
&\overset{\text{B}}{\rightsquigarrow}_{\mathsf{t}} \mathsf{coh}\,(D^n : d_n \to_{S^{n-1}} d_n)[t] \\
&\overset{\text{D}}{\rightsquigarrow} t
\end{aligned}$$

Here, $n := \mathsf{dim}(t)$. As a consequence, we prove by induction on subterms that $N(R(t)) = N(t)$:

$$R(t) \equiv R(\mathsf{coh}\,(\Gamma : U)[\sigma])$$

$$\equiv P\big(\mathsf{coh}\,(\Gamma : R(U))[R(\sigma)]\big)$$
$$= \mathsf{coh}\,(\Gamma : R(U))[R(\sigma)]$$
$$= \mathsf{coh}\,(\Gamma : U)[\sigma]$$

This completes the proof. □

References

[1] Thorsten Altenkirch and Ambrus Kaposi. Type theory in type theory using quotient inductive types. *ACM SIGPLAN Notices*, 51(1):18–29, 2016.

[2] Dimitri Ara. Sur les ∞-groupoides de Grothendieck et une variante ∞-catégorique. 2010. Ph.D. dissertation, Université Paris Diderot.

[3] John Baez and Mike Stay. Physics, topology, logic and computation: a rosetta stone. In *New Structures for Physics*, pages 95–172. Springer, 2010.

[4] Michael Batanin. Monoidal globular categories as a natural environment for the theory of weak n-categories. *Adv. Math.*, 136(1):39–103, 1998.

[5] Michael Batanin, Denis-Charles Cisinski, and Mark Weber. Multitensor lifting and strictly unital higher category theory. *Theory and Applications of Categories*, 28:804–856, 2013.

[6] Thibaut Benjamin. Type theory for weak ω-categories. 2020. Ph.D. dissertation.

[7] John Michael Boardman and Rainer M Vogt. *Homotopy invariant algebraic structures on topological spaces*, volume 347. Springer, 2006.

[8] Guillaume Brunerie. On the homotopy groups of spheres in homotopy type theory. *arXiv preprint arXiv:1606.05916*, 2016.

[9] Alexandre Buisse and Peter Dybjer. Towards formalizing categorical models of type theory in type theory. *Electronic Notes in Theoretical Computer Science*, 196:137–151, 2008.

[10] James Chapman. Type theory should eat itself. *Electronic Notes in Theoretical Computer Science*, 228:21–36, 2009.

[11] Nils Anders Danielsson. A formalisation of a dependently typed language as an inductive-recursive family. In *International Workshop on Types for Proofs and Programs*, pages 93–109. Springer, 2006.

[12] Christoph Dorn. Associative n-categories. 2018. Ph.D. dissertation. arXiv:1812.10586.

[13] Eric Finster and Samuel Mimram. A type-theoretical definition of weak ω-categories. In *Proceedings of LICS 2017*, 2017. arXiv:1706.02866.

[14] R. Gordon, A. J. Power, and Ross Street. Coherence for tricategories. *Mem. Amer. Math. Soc.*, 117(558):vi+81, 1995.

[15] Barnaby P Hilken. Towards a proof theory of rewriting: the simply typed 2λ-calculus. *Theoretical Computer Science*, 170(1-2):407–444, 1996.

[16] Dexter C Kozen. *Automata and computability.* Springer Science & Business Media, 2012.

[17] Tom Leinster. *Higher operads, higher categories*, volume 298. Cambridge University Press, 2004.

[18] Jacob Lurie. *Higher topos theory.* Princeton University Press, 2009.

[19] Georges Maltsiniotis. Grothendieck ∞-groupoids, and still another definition of ∞-categories. 2010. arXiv preprint arXiv:1009.2331.

[20] David Reutter and Jamie Vicary. High-level methods for homotopy construction in associative n-categories. In *Proceedings of LICS 2019*, 2019. `arXiv:1902.03831`.

[21] Philip Saville. *Cartesian closed bicategories: type theory and coherence.* PhD thesis, Ph. D. thesis, University of Cambridge (Submitted), 2019.

[22] Robert AG Seely. Modelling computations: a 2-categorical framework. In *LICS*, pages 65–71, 1987.

[23] Michael Shulman. Homotopy type theory should eat itself (but so far, it's too big to swallow). *Homotopy type theory blog post*, 2014. .

[24] The Univalent Foundations Program. *Homotopy Type Theory: Univalent Foundations of Mathematics.* `https://homotopytypetheory.org/book`, Institute for Advanced Study, 2013.