ON CLASSICAL DETERMINATE TRUTH

LUCA CASTALDO
Munich Center for Mathematical Philosophy, LMU Munich
and

CARLO NICOLAI
King's College London

Abstract. The paper proposes and studies new classical, type-free theories of truth and determinateness with unprecedented features. The theories are fully compositional, strongly classical (namely, their internal and external logics are both classical), and feature a *defined* determinateness predicate satisfying desirable and widely agreed principles. The theories capture a conception of truth and determinateness according to which the generalizing power associated with the classicality and full compositionality of truth is combined with the identification of a natural class of sentences—the determinate ones—for which clear-cut semantic rules are available. Our theories can also be seen as the *classical closures* of Kripke–Feferman truth: their ω-models, which we precisely pin down, result from including in the extension of the truth predicate the sentences that are satisfied by a Kripkean closed-off fixed-point model. The theories compare to recent theories proposed by Fujimoto and Halbach, featuring a primitive determinateness predicate. In the paper we show that our theories entail all principles of Fujimoto and Halbach's theories, and are proof-theoretically equivalent to Fujimoto and Halbach's CD^+. We also show establish some negative results on Fujimoto and Halbach's theories: such results show that, unlike what happens in our theories, the primitive determinateness predicate prevents one from establishing clear and unrestricted semantic rules for the language with type-free truth.

§1. A conception of truth (and determinateness).

In this work we offer a cluster of formal theories with unprecedented features and prove several results about them. The theories are accompanied by a conception of truth (and determinateness), which we outline in this introductory section.[1]

1.1. Desiderata for truth. Following a long-standing tradition in formal theories of truth, starting at least with [17], we believe that truth should be *type-free*; the truth predicate (provably) applies to sentences containing itself.

A type-free notion of truth, as it is well-known, calls for an account of the Liar and related paradoxes. We will shortly outline the solution implicit in our theories.

Received: September 6, 2024.
2020 *Mathematics Subject Classification*: Primary 03A99, 03B50, 03F03, 03F25.
Key words and phrases: formal theories of truth, axiomatic theories of truth, determinateness, liar paradox, fixed-point semantics.

[1] A full exposition and philosophical defence of the conception is carried out in a companion paper.

doi:10.1017/S1755020325100968

Independently of the details, though, the view we put forward is *strongly classical*: not only the external logic of the theories should be classical, but we require the logic of their internal theories[2] to be classical as well. Our reasons for countenancing a classical external logic are the familiar ones. By endorsing a nonclassical theory of truth, one is bound to sever the links between the theory of truth and classical mathematics. This is a symptom of what McGee called 'degradation of methodology' [18, chap. 4]; the standards in one's truth-theoretic theorizing shouldn't be any lower than the ones employed in theorizing about core scientific subjects.[3]

But, we argued, the logic of the internal theory should be classical as well. The truth predicate, in natural and formal languages, is used to express important general claims. To perform its generalization role in full, the truth predicate should be *fully compositional*, namely, commuting with all classical logical connectives, and without any type restrictions. Full compositionality is for instance required to establish the truth of the laws of our classical external logic, such as 'all instances of the law of excluded middle are true' [22]. Compositionality is also essential to capture logical inferences involving arbitrary sentences (or propositions), what [12] calls *blind deductions*. For instance, to directly formalize the argument

> Everything that Gerhard says about cut-elimination is true. David asserted the negation of some of Gerhard's claims. Therefore, something David asserted is not true.

one requires commutation of truth with negation in fully quantified form (as well as no type restriction).

1.2. From significance to semantic (in)sensitivity. A type-free, strongly compositional notion of truth requires a principled account of how the truth predicate can self-apply *sine contradictione*. The naive principle

$$\text{for any } \varphi, \text{`}\varphi \text{ is true' is true iff } \varphi \text{ is true} \tag{1}$$

is in fact inconsistent with the assumptions above. In their recent paper [13], Kentaro Fujimoto and Volker Halbach propose to restrict the principle (1) to what they call *determinate* sentences, while retaining a fully compositional truth predicate. The proposed restriction of (1), in turn, entails the restricted T-schema

$$\mathrm{D}\ulcorner A \urcorner \rightarrow (\mathrm{T}\ulcorner A \urcorner \leftrightarrow A), \tag{2}$$

where D is a determinateness predicate.

Several theories satisfying (2) can be found in the context of Kripke–Feferman truth. Well-known examples are Feferman's KF and DT—from [5, 6], respectively. Such theories rely on a Russellian notion of range of significance for the interpretation of D, which in the context of Feferman's theories can be further analyzed as the sentences which *possess a classical truth value* relative to the intended semantics:

> I agree with Russell (1908) that every predicate has a domain of significance, and it makes sense to apply the predicate only to objects in that domain. In the case of truth, that domain D consists of the sentences that are meaningful and determinate, that is, have a definite

[2] For a theory T, its internal theory is defined as $\{A \mid \text{`}\ulcorner A \urcorner \text{ is true' is provable in } T\}$.
[3] For more on the mathematical costs of adopting a nonclassical logic, see [8, 10, 15].

> truth value, true or false. D includes various but not necessarily all grammatically correct sentences that involve the notion of truth itself [6, p. 206].

A similar picture is endorsed by Reinhardt, in [23, 24], who labels classically true or false sentences 'significant'. A well-known feature of the restriction of the T-schema to determinate sentences, already extensively discussed in [24],[4] is that any such restriction would entail the existence of sentences that are provable and yet (provably) not determinate.

Fujimoto and Halbach, however, reject the Russellian view and claim that truth can be (provably) applied to sentences that are not determinate. They suggest moving from the notions of range of significance to a notion of *semantic sensitivity*: 'some sentences, including liar sentences, are sensitive to the addition of another layer of truth. Stacking an additional layer of truth onto the liar sentence will change its semantic status; but that does not mean that the truth predicate cannot meaningfully be applied to it' [13, p. 256]. The shift from 'significance' to 'sensitivity' has the effect of weakening the Reinhardt–Bacon issue; it is not problematic to prove sentences that are (provably) semantically sensitive.

The shift from significance to semantic sensitivity is, in our view, a valuable step forward. Fujimoto and Halbach, however, only devote a few remarks to the notion. We sketch below our own interpretation of semantic sensitivity in the context of type-free truth, and employ it to highlight the benefits of the theories proposed in the paper.

Another distinctive feature of Fujimoto and Halbach's account is that determinateness cannot be explained in terms of truth. Their theories feature a *primitive* determinateness predicate as well as a primitive truth predicate. Fujimoto and Halbach suggest that this has to be so; in discussing Feferman's notion of determinateness from [6], they write '[Feferman] took D as definable in terms of truth by stipulating $\mathrm{D}x :\leftrightarrow \mathrm{T}x \vee \mathrm{F}x$; but this definition does not yield the desired properties of D in our theory and we consequently introduce determinateness as primitive notion' [13, p. 257].

As some of our results will indicate, the primitive nature of the determinateness predicate significantly complicates the semantic analysis of the truth-theoretic language. Fortunately, as we will show, virtually all of the theoretical benefits sought for by Fujimoto and Halbach can be met by defining determinateness in terms of truth.

1.3. Logic and semantics. A formal theory of type-free truth plays, in our account, a double role. It provides general laws for truth that are used in several theoretical contexts, including semantic theorizing. As a result of the discussion above, these must include compositional axioms in combination with classical (external and internal) logic. In addition, the theory of truth provides a formalization of clear semantic rules to be applied to a well-defined and sufficiently comprehensive class of sentences of the (type-free) truth-theoretic language. Because of paradox, such rules may not coincide with the classical semantic rules. We will see that this is precisely what the theories introduced below will be able to achieve.

The semantic rules that we identify involve decomposition of semantic values according to the familiar Strong-Kleene truth conditions, including full disquotation

[4] See also [1]. One applies basic logical steps to a sentence 'saying of itself' that it's either not determinate or not true.

for the truth predicate. In the class of models we privilege, the natural collection of sentences to which these rules unrestrictedly apply are the Kripke-determinate ones, that is sentences that are in the extension of the relevant consistent fixed-point model in the sense of [17].[5] Relative to each such model, these are the *determinate* sentences, and the T-schema holds unrestrictedly for them. Relative to the minimal fixed point, the determinate sentences are simply the grounded ones.

As a consequence, just like in Fujimoto and Halbach's picture, the determinate sentences are semantically insensitive in the sense that their semantic status is not affected by one or more iteration of the truth predicate on the sentence. However, unlike what happens in Fujimoto and Halbach's theory, we can provide fully general and uniform semantic rules for the semantic analysis of the language with a type-free truth predicate. As the results of §3 will establish, such clearly defined rules cannot be found in Fujimoto and Halbach's approach.

The logical role of truth is fulfilled, in our framework, primarily by full compositional principles. These cannot be semantically insensitive, as they are quantified sentences with instances that cannot be determinate. In fact, one of the novel features of our approach is the uniform combination of Kripkean fixed-point semantics with full compositionality. This is achieved by a move that we call *classical closure of Kripkean truth*. Formally, this prescribes that the extension of the truth predicate contain the sentences that are satisfied by the relevant closed-off fixed-point model. Conceptually, it demands that the truth-theorist engaged in semantic theorizing fully embrace the generalizing power afforded by the truth predicate.

It is worth noting that, as a direct consequence of full compositionality, some sentences happen to be in the extension of the classical closure of Kripkean truth only in virtue of the generalizing role of truth. This is for instance the case of classical logical truths. Take, e.g., the sentence $\lambda \vee \neg\lambda$ for λ a liar sentence. The sentence is true, but any further iteration of truth on it cannot be.

Ultimately, we are focusing on the classical models resulting from taking the classical closure of consistent fixed points as extension of the truth predicate. Given the set up, full compositional axioms will be satisfied in such models. Moreover, since the determinate sentences can be readily defined as the ones that are recognized as true or false (i.e., have a true negation) in the classical closure of Kripkean truth, a sentence will be determinate precisely when the sentence stating that it is true or false is itself true in the model. The model thus provides a strongly classical environment for the study of a traditional notion of determinateness.

1.4. Axiomatization and $\mathbb{N}$-categoricity. The paper is primarily concerned with axiomatic theories of truth capturing the intended semantics just sketched. Without axiomatization, semantic constructions are typically hard to pin down. Especially in the context of classical theories, a robust sense in which an axiomatic theory captures a semantic construction is given by the notion of $\mathbb{N}$-categoricity introduced by [9]. Let $\mathscr{E} \subseteq \mathscr{P}(\omega)$ be a collection of extensions of the truth predicate; a theory T $\mathbb{N}$-categorically axiomatizes $\mathscr{E}$ just in case

$$(\mathbb{N}, S) \vDash T \text{ iff } S \in \mathscr{E}.$$

Another remarkable feature of the theories we introduce in the paper is that they turn out to be $\mathbb{N}$-categorical with respect to the models described in the previous section.

[5] However, we will also consider classes of models with inconsistent extensions.

Specifically, a structure $(\mathbb{N}, S)$ models our theories if and only if S is the classical closure of a consistent Kripkean fixed point. The $\mathbb{N}$-categoricity of our theories marks another key difference with Fujimoto and Halbach's approach afforded by the defined determinateness predicate. As the results below will indicate, while some models of Fujimoto and Halbach's theories are akin to the classical closure of Kripkean truth, some others are not.

Therefore, there is a precise sense in which the conception of truth outlined in this section—and embodied in the class of ω-models just described—corresponds to the laws of truth we propose and study in the paper.

1.5. Summary of the main results. From the technical side, we study theories that are closely related to the ones studied by Fujimoto and Halbach in their [13], but with important differences that will be highlighted in due course. Fujimoto and Halbach mainly focus on two systems, CD and CD^+ (see Definition 1). CD^+ is obtained by strengthening the axiom

$$\forall t(\mathrm{D}t^\circ \to \mathrm{TD}t) \tag{T2}$$

of CD with the (plausible) biconditional

$$\forall t(\mathrm{D}t^\circ \leftrightarrow \mathrm{TD}t). \tag{T2$^+$}$$

As we will explain shortly, in the paper we focus exclusively on CD^+. Here is a summary of the main results.

(a) We introduce a theory of truth, $\mathsf{CD}^+_{\mathrm{T}}$, whose axioms are the axioms of CD^+ except that the determinateness predicate $\mathrm{D}x$ is defined as $\mathrm{TT}x \vee \mathrm{TF}x$ (Definition 7). We show that the theory is consistent by displaying a natural class of ω-models for it (Proposition 10), and that the theory is mutually reducible, in a strong sense, with CD^+.

(b) We show that $\mathsf{CD}^+_{\mathrm{T}}$ can be reaxiomatized as what we will call KF's classical closure ($\mathsf{CKF}_{\mathsf{cs}}$, Definition 15), i.e., the fully compositional theory of type-free truth (with no axioms for determinateness) obtained by axiomatizing $\mathsf{KF} + \mathsf{CONS}$ within the scope of a truth predicate satisfying unrestricted compositional principles (Proposition 25). A particularly nice feature of $\mathsf{CKF}_{\mathsf{cs}}$ (and hence of $\mathsf{CD}^+_{\mathrm{T}}$) is its $\mathbb{N}$-categoricity with respect to the classical closures of consistent Kripkean fixed points (Proposition 21). This feature is not available for CD^+.

(c) While studying $\mathsf{CD}^+_{\mathrm{T}}$ and $\mathsf{CKF}_{\mathsf{cs}}$, we provide additional results on *both theories* CD and CD^+: we show that they cannot prove most of the key axioms of $\mathsf{CKF}_{\mathsf{cs}}$; some of these axioms were proved to be conservative over CD^+ in [13]. Our results show that the behaviour of CD and CD^+'s truth predicate, while fully classical and compositional on the surface, becomes highly irregular within two or more layers of truth.

(d) We show that our results are stable under a dual definition of $\mathrm{D}x$, which is suitable for a complete (but not consistent) truth predicate. Specifically, we introduce the theory $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$, whose axioms are those of CD^+ except that $\mathrm{D}x$ is defined as $\neg\mathrm{TT}x \vee \neg\mathrm{TF}x$, and show that it is mutually reducible with CD^+. We also show that $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$ can be reaxiomatized as the classical closure of $\mathsf{KF} + \mathsf{COMP}$.

The reason for focusing on CD^+ is that the natural class of models referred to in (a) and the axiomatizations mentioned in (b)–(d) license all principles of CD^+.[6]

1.6. Technical preliminaries. We work over the language of arithmetic $\mathcal{L}_{\mathbb{N}}$, which extends the standard signature $\{\overline{0}, \mathrm{S}, +, \times\}$ with finitely many function symbols for elementary syntactic operations (see the next paragraph for more details). Let $\mathcal{L}_{\mathrm{T}} := \mathcal{L}_{\mathbb{N}} \cup \{\mathrm{T}\}$ and $\mathcal{L}_{\mathrm{D}} := \mathcal{L}_{\mathrm{T}} \cup \{\mathrm{D}\}$, where T and D are unary truth and determinateness predicates, respectively. We fix a canonical Gödel numbering of $\mathcal{L}_{\mathrm{D}}$-expressions and a formalization of syntactic notions and operations as it can be found, for instance, in [14]. Following standard practice, we take Peano arithmetic to be the first-order theory in which this formalization is carried out, although of course much weaker systems would suffice. We write PAT for the theory obtained by formulating the axiom of PA in the expanded language $\mathcal{L}_{\mathrm{T}}$.

The additional function symbols of $\mathcal{L}_{\mathbb{N}}$ include a symbol for the standard numeral function $\mathrm{num}(x)$ sending a number to the code of its numeral. A standard dot-notation to denote such symbols:

Operation	Function in $\mathcal{L}_{\mathbb{N}}$
$\#t, \#s \mapsto \#(t = s)$	$\dot{=}$
$\#\varphi \mapsto \#(\neg\varphi)$	$\dot{\neg}$
$\#\varphi, \#\psi \mapsto \#(\varphi \wedge \psi)$	$\dot{\wedge}$
$\#v_k, \#\varphi \mapsto \#(\forall v_k \varphi)$	$\dot{\forall}$
$\#t \mapsto \#\mathrm{T}(t)$	$\dot{\mathrm{T}}$
$\#t \mapsto \#\mathrm{D}(t)$	$\dot{\mathrm{D}}$.

Using the conventions above, we define falsity as true negation, that is, $\mathrm{F}x :\leftrightarrow \mathrm{T}\dot{\neg}x$. We also take the following $\mathcal{L}_{\mathbb{N}}$-predicates to abbreviate the equations for the (elementary) characteristic function for such sets:

$\mathrm{Term}(x)$ $(\mathrm{Cterm}(x)) := x$ is the Gödel number of a (closed) term;
$\mathrm{Fml}^n_{\mathcal{L}}(x)$ $(\mathrm{Sent}_{\mathcal{L}}(x)) := x$ is the Gödel number of an $\mathcal{L}$-formula with at most n (0) free distinct variables.

As in Halbach's monograph [16], we will employ a functional notation x° abbreviating the formula representing in $\mathcal{L}_{\mathbb{N}}$ the evaluation function for closed terms of $\mathcal{L}_{\mathbb{N}}$ and is such that $\ulcorner t \urcorner^\circ = t$ for closed terms t.

Roman uppercase letters $A, B, C, \ldots$ range over formulae of $\mathcal{L}_{\mathrm{D}}$. Greek lowercase letters $\varphi, \psi, \xi, \ldots$ will be used to abbreviate quantification over formal sentences, while $\varphi(v), \psi(v), \xi(v), \ldots$ will be used to abbreviate quantification over formal formulae with at most v free. Also, for the sake of readability, we suppress codes, dots, and Quine corners when there is no danger of confusion. Similar conventions apply to $\mathcal{L}_{\mathrm{T}}$. The expression $e(t/v_k)$, sometimes abbreviated as $e(t)$, represents the syntactic substitution of a term t for a variable v_k in an expression e. So, for example, the expressions

[6] It is unclear whether a strategy of the kind we employ to define the truth predicate of CD^+ is available for CD.

$$(\forall \varphi, \psi : \mathcal{L}_{\mathrm{D}})(\mathrm{T}(\varphi \wedge \psi) \leftrightarrow \mathrm{T}\varphi \wedge \mathrm{T}\psi)$$
$$\forall t(\mathrm{TT}t \to \mathrm{T}t^\circ)$$
$$(\forall \varphi(v) : \mathcal{L}_{\mathrm{D}})(\mathrm{T}(\forall v\varphi) \leftrightarrow \forall t\, \mathrm{T}\varphi(t/v))$$

are short for, respectively,

$$\forall x \forall y(\mathrm{Sent}_{\mathcal{L}_{\mathrm{D}}}(x) \wedge \mathrm{Sent}_{\mathcal{L}_{\mathrm{D}}}(y) \to (\mathrm{T}(x \underset{\cdot}{\wedge} y) \leftrightarrow \mathrm{T}x \wedge \mathrm{T}y))$$
$$\forall x(\mathrm{Cterm}(x) \to \mathrm{T}\underset{\cdot}{\mathrm{T}}x \to \mathrm{T}x^\circ).$$
$$\forall x \forall y\big(\mathrm{Fml}^1_{\mathcal{L}_{\mathrm{D}}}(x) \wedge \mathrm{Var}(y) \to (\mathrm{T}\underset{\cdot}{\forall} yx \leftrightarrow \forall z(\mathrm{Cterm}(z) \to \mathrm{T}x(z)))\big).$$

We will also write $A \in X$ instead of $\#A \in X$ or, for t ranging over closed terms, $\mathrm{TTT}t$ instead of $\mathrm{T}(\underset{\cdot}{\mathrm{T}}\mathrm{num}\underset{\cdot}{\mathrm{T}}t)$, etc.

We use standard notions of *relative translation* and *relative interpretation* as it can be found, for instance, in [16, 27]. We use the notions of $\mathcal{L}_{\mathbb{N}}$-*translations* and $\mathcal{L}_{\mathbb{N}}$-*interpretations*: an $\mathcal{L}_{\mathbb{N}}$-translation is a translation between theories in $\mathcal{L}_1, \mathcal{L} \supseteq \mathcal{L}_{\mathbb{N}}$ which does not relativize quantifiers and collapses into the identity function when restricted to $\mathcal{L}_{\mathbb{N}}$. An $\mathcal{L}_{\mathbb{N}}$-interpretation is an $\mathcal{L}_{\mathbb{N}}$-translation that preserves provability in the standard way. Indeed, the special case of $\mathcal{L}_{\mathbb{N}}$-interpretability in which $\mathcal{L}_1, \mathcal{L} \supseteq \mathcal{L}_{\mathbb{N}}$ expand the arithmetical signature with truth predicates is the notion of truth definability in the sense of [11].

§2. CD^+.

Fujimoto and Halbach introduce the theory CD^+ in [13].

Definition 1. CD^+ *consists of the following extension of* PA *in the language* $\mathcal{L}_{\mathrm{D}}$ (*where* D *and* T *are allowed to appear in induction*):

$$\forall s \forall t(\mathrm{T}(s = t) \leftrightarrow s^\circ = t^\circ) \tag{T1}$$
$$\forall t(\mathrm{D}t^\circ \leftrightarrow \mathrm{TD}t) \tag{T2$^+$}$$
$$\forall t(\mathrm{D}t^\circ \to (\mathrm{TT}t \leftrightarrow \mathrm{T}t^\circ)) \tag{T3}$$
$$(\forall \varphi : \mathcal{L}_{\mathrm{D}})(\mathrm{T}(\neg\varphi) \leftrightarrow \neg\mathrm{T}\varphi) \tag{T4}$$
$$(\forall \varphi, \psi : \mathcal{L}_{\mathrm{D}})(\mathrm{T}(\varphi \wedge \psi) \leftrightarrow \mathrm{T}\varphi \wedge \mathrm{T}\psi) \tag{T5}$$
$$(\forall \varphi(v) : \mathcal{L}_{\mathrm{D}})(\mathrm{T}(\forall v\varphi) \leftrightarrow \forall t\, \mathrm{T}\varphi(t/v)) \tag{T6}$$
$$\forall s \forall t\, \mathrm{D}(s = t) \tag{D1}$$
$$\forall t(\mathrm{DT}t \leftrightarrow \mathrm{D}t^\circ) \tag{D2}$$
$$\forall t(\mathrm{DD}t \leftrightarrow \mathrm{D}t^\circ) \tag{D3}$$
$$(\forall \varphi : \mathcal{L}_{\mathrm{D}})(\mathrm{D}(\neg\varphi) \leftrightarrow \mathrm{D}\varphi) \tag{D4}$$
$$(\forall \varphi, \psi : \mathcal{L}_{\mathrm{D}})\big(\mathrm{D}(\varphi \wedge \psi) \leftrightarrow ((\mathrm{D}\varphi \wedge \mathrm{D}\psi) \vee (\mathrm{D}\varphi \wedge \mathrm{T}\neg\varphi) \vee (\mathrm{D}\psi \wedge \mathrm{T}\neg\psi))\big) \tag{D5}$$
$$(\forall \varphi(v) : \mathcal{L}_{\mathrm{D}})(\mathrm{D}(\forall v\varphi) \leftrightarrow (\forall t\, \mathrm{D}\varphi(t/v) \vee \exists t\, \mathrm{D}\varphi(t/v) \wedge \mathrm{T}\neg\varphi(t/v)). \tag{D6}$$
$$(\forall \varphi(v) : \mathcal{L}_{\mathrm{D}})\forall s \forall t\big(s^\circ = t^\circ \to (\mathrm{T}\varphi(s) \leftrightarrow \mathrm{T}\varphi(t))\big) \tag{R1}$$
$$(\forall \varphi(v) : \mathcal{L}_{\mathrm{D}})\forall s \forall t\big(s^\circ = t^\circ \to (\mathrm{D}\varphi(s) \leftrightarrow \mathrm{D}\varphi(t))\big). \tag{R2}$$

Fujimoto and Halbach establish the consistency of CD^+—and, of course, of CD—by exhibiting an ω-model. Since the construction will be relevant later on, we repeat it here. Let $\mathcal{D}(x)$ be the positive inductive definition associated with the (right-to-left direction) of the D-axioms D1–D6 [13, sec. 4]. Let

$$\Gamma_Y(X) = \{n \mid (\mathbb{N}, X, Y) \vDash \mathcal{D}(\overline{n})\},$$

and define

$$\begin{aligned} D_0 &:= \varnothing & T_0 &:= \varnothing \\ D_{\alpha+1} &:= \Gamma_{T_\alpha}(D_\alpha) & T_{\alpha+1} &:= \{A \in \mathcal{L}_{\mathrm{D}} \mid (\mathbb{N}, D_\alpha, T_\alpha) \vDash A\} \\ D_\lambda &:= \bigcup_{\beta<\lambda} D_\beta & T_\lambda &:= \bigcup_{\beta<\lambda} (D_\beta \cap T_\beta). \end{aligned}$$

Then D_{ω_1} is a fixed point of $\Gamma_{T_{\omega_1}}$. Let us denote D_{ω_1} and T_{ω_1} by D_∞ and T_∞, respectively. Let:

$$\mathbb{T}_\infty = \{A \in \mathcal{L}_{\mathrm{D}} \mid (\mathbb{N}, D_\infty, T_\infty) \vDash A\}.$$

Lemma 2 [13, theorem 4.8]. $(\mathbb{N}, D_\infty, \mathbb{T}_\infty) \vDash \mathsf{CD}^+$.

The proof-theoretic analysis of CD^+ is quite straightforward. It reduces the system to a theory resulting from a combination of typed truth with standard Kripke–Feferman systems. It employs CT[KF + CONS], which is the result of enriching the Kripke–Feferman theory with consistency with a *typed, Tarskian* truth predicate $\mathbf{T}$ on top, but a version of KF with the completeness axiom would work as well.

Definition 3. KF *is the system in* $\mathcal{L}_{\mathrm{T}}$ *extending* PAT *with the following axioms*:

$$\forall s \forall t (\mathrm{T}(s = t) \leftrightarrow s^\circ = t^\circ) \wedge \forall s \forall t (\mathrm{T}(s \neq t) \leftrightarrow s^\circ \neq t^\circ) \tag{KF1}$$

$$\forall t (\mathrm{TT}t \leftrightarrow \mathrm{T}t^\circ) \wedge \forall t (\mathrm{FT}t \leftrightarrow \mathrm{F}t^\circ \vee \neg \mathrm{Sent}(t^\circ)) \tag{KF2}$$

$$(\forall \varphi\colon \mathcal{L}_{\mathrm{T}})(\mathrm{F}\neg\varphi \leftrightarrow \mathrm{T}\varphi) \tag{KF3}$$

$$(\forall \varphi, \psi\colon \mathcal{L}_{\mathrm{T}})((\mathrm{T}(\varphi \wedge \psi) \leftrightarrow \mathrm{T}\varphi \wedge \mathrm{T}\psi) \wedge (\mathrm{F}(\varphi \wedge \psi) \leftrightarrow \mathrm{F}\varphi \vee \mathrm{F}\psi)) \tag{KF4}$$

$$(\forall \varphi(v)\colon \mathcal{L}_{\mathrm{T}})\big((\mathrm{T}(\forall v \varphi) \leftrightarrow \forall t\, \mathrm{T}\varphi(t/v)) \wedge (\mathrm{T}(\exists v \varphi) \leftrightarrow \exists t\, \mathrm{T}\varphi(t/v))\big). \tag{KF5}$$

KF + CONS *is obtained by adding to* KF *the axiom*

$$(\forall \varphi\colon \mathcal{L}_{\mathrm{T}})(\mathrm{F}\varphi \rightarrow \neg \mathrm{T}\varphi); \tag{Cons}$$

KF + COMP *is obtained by adding to* KF *the converse claim*

$$(\forall \varphi\colon \mathcal{L}_{\mathrm{T}})(\neg \mathrm{T}\varphi \rightarrow \mathrm{F}\varphi). \tag{Comp}$$

Definition 4 (CT[KF + CONS]). *We expand* $\mathcal{L}_{\mathrm{T}}$ *with an additional truth predicate* $\mathbf{T}$ *and call the resulting language* $\mathcal{L}_{\mathrm{T},\mathbf{T}}$. CT[KF + CONS] *is the theory in* $\mathcal{L}_{\mathrm{T},\mathbf{T}}$ *extending* KF + CONS *with the following axioms*:

$$(\mathbf{T}(s = t) \leftrightarrow s^\circ = t^\circ) \wedge \mathbf{T}\mathrm{T}t \leftrightarrow \mathrm{T}t^\circ \tag{T1}$$

$$(\forall \varphi\colon \mathcal{L}_{\mathrm{T}})(\mathbf{T}(\neg\varphi) \leftrightarrow \neg \mathbf{T}\varphi) \tag{T2}$$

$$(\forall \varphi, \psi\colon \mathcal{L}_{\mathrm{T}})(\mathbf{T}(\varphi \wedge \psi) \leftrightarrow \mathbf{T}\varphi \wedge \mathbf{T}\psi) \tag{T3}$$

$$(\forall \varphi(v)\colon \mathcal{L}_{\mathrm{T}})(\mathbf{T}(\forall v \varphi) \leftrightarrow \forall t\, \mathbf{T}(\varphi(t/v))). \tag{T4}$$

We assume a standard notation for ordinals below the Feferman–Schütte ordinal Γ_0 (see, e.g., [20, chap. 2]). In particular, we denote with ε_α the (code of) the α^{th} fixed point of the function $\lambda x.\, \omega^x$.

LEMMA 5 [13, theorem 7.12]. CD^+ *is mutually* $\mathcal{L}_{\mathbb{N}}$*-interpretable with* $\mathsf{CT}[\mathsf{KF}+\mathsf{CONS}]$, *and therefore with ramified truth/analysis up to any* $\alpha < \varepsilon_{\varepsilon_0}$.

Proof Idea. To interpret CD^+ in $\mathsf{CT}[\mathsf{KF}+\mathsf{CONS}]$, external occurrences of T are translated by the Tarskian **T**, whereas internal occurrences of T are translated homophonically as T. D becomes $\mathrm{T} \cup \mathrm{F}$ (the sentences with 'a classical truth value' in $\mathsf{KF}+\mathsf{CONS}$). Conversely, to interpret $\mathsf{CT}[\mathsf{KF}+\mathsf{CONS}]$ in CD^+, the Tarskian **T** is translated as T, and the $\mathsf{KF}+\mathsf{CONS}$ truth predicate T is instead translated as the 'determinate truths' of CD^+, i.e., $\mathrm{T} \cap \mathrm{D}$. □

§3. $\mathsf{CD}^+_{\mathrm{T}}$ and its intended models. Our first result consists in displaying a natural class of models for type-free truth, the *classical closures of (consistent) Kripkean fixed points*. We will also show that such structures model the axioms of Fujimoto and Halbach's CD^+. Crucially, in these models, D is defined in terms of the notion of Kripkean determinateness. In the next section, we will provide direct axiomatizations of such structures in the language $\mathcal{L}_{\mathrm{T}}$.

We recall the arithmetical operator associated with the Strong-Kleene version of the fixed-point semantics.

DEFINITION 6 (K-jump). *For any* $X \subseteq \mathrm{Sent}_{\mathcal{L}_{\mathrm{T}}}$,

$$\begin{aligned} n \in \mathscr{K}(X) :\leftrightarrow n \in \mathrm{Sent}_{\mathcal{L}_{\mathrm{T}}} \wedge & \\ & \big(\exists s \exists t (n = (s = t) \wedge s^\circ = t^\circ) \vee \\ & \exists s \exists t (n = (s \neq t) \wedge s^\circ \neq t^\circ) \vee \\ & \exists t (n = (\mathrm{T}t) \wedge t^\circ \in X) \vee \\ & \exists t (n = (\neg \mathrm{T}t) \wedge \neg t^\circ \in X \vee t^\circ \in \omega \setminus \mathrm{Sent}_{\mathcal{L}_{\mathrm{T}}}) \vee \\ & \exists \varphi (n = \neg\neg\varphi \wedge \varphi \in X) \vee \\ & \exists \varphi, \psi (n = (\varphi \wedge \psi) \wedge \varphi \in X \wedge \psi \in X) \vee \\ & \exists \varphi, \psi (n = \neg(\varphi \wedge \psi) \wedge \neg\varphi \in X \vee \neg\psi \in X) \\ & \exists \varphi(v) (n = (\forall v \varphi) \wedge \forall t\, \varphi(t/v) \in X) \\ & \exists \varphi(v) (n = \neg \forall v \varphi \wedge \exists t\, \neg\varphi(t/v) \in X)\big). \end{aligned}$$

X occurs positively in $\mathscr{K}(X)$ and the K-jump is monotone. Therefore, it will have fixed points. Note that fixed points of $\mathscr{K}$ are closed under Strong-Kleene logic and are such that $(\neg)\varphi \in X$ iff $(\neg)\mathrm{T}\varphi \in X$.[7]

In this section and the next we are interested in *consistent fixed points of* $\mathscr{K}$ (and theories that are sound with respect to them), that is, sets X such that $X = \mathscr{K}(X)$ and such that there's no φ such that $\varphi \wedge \neg\varphi \in X$; we will consider other fixed points in a later section. The reason for this is both technical and conceptual. Technically, consistency delivers a simpler definition of determinateness and provides a basis for results about complete models and theories presented in later sections. Conceptually, we believe that a consistent extension of truth provides a more attractive notion of determinateness, as well as being in continuity with the above-mentioned works on truth-theoretic determinateness by Kripke, Reinhardt and Feferman.

[7] For more details on fixed-point semantics, we refer to [16, sec. 15.1] and [18, secs. 4 and 5].

For a consistent fixed-point X, let

$$\mathbb{T}_X = \{A \in \mathcal{L}_\mathrm{T} \mid (\mathbb{N}, X) \vDash A\}. \tag{3}$$

Incidentally, this set is considered by Fujimoto and Halbach as a model of $\mathsf{CT}[\mathsf{KF}+\mathsf{CONS}]$ [13, p. 251]. We will now show that this set can be used to provide a direct model construction for the principles of CD^+, and not only for a typed theory of truth interpreting CD^+.

DEFINITION 7 (CD^+_T). CD^+_T *is the theory in* $\mathcal{L}_\mathrm{T}$ *whose axioms are the axioms of* CD^+, *but where D is now* defined *in terms of* T *as*

$$\mathrm{D}x :\leftrightarrow \mathrm{T}\dot{\mathrm{T}}\mathrm{num}(x) \vee \mathrm{T}\dot{\mathrm{F}}\mathrm{num}(x) \quad (\text{abbr. } \mathrm{TT}x \vee \mathrm{TF}x). \tag{4}$$

CD^+_T, just like CD^+, delivers the intended restriction to the T-schema to determinate sentences.

OBSERVATION 8. *For any* $A \in \mathcal{L}_\mathrm{T}$: $\mathsf{CD}^+_\mathrm{T} \vdash \mathrm{D}\ulcorner A \urcorner \to (\mathrm{T}\ulcorner A \urcorner \leftrightarrow A)$.

Proof. By external induction on the complexity of A. Note that the case in which A is $\mathrm{T}t$ is an axiom of CD^+_T. $\square$

Our next goal is to show that, given a consistent fixed-point X, the structure $(\mathbb{N}, \mathbb{T}_X)$ is a model of CD^+_T. Since $(\mathbb{N}, X) \vDash \mathsf{KF}+\mathsf{CONS}$, and since $\mathsf{KF}+\mathsf{CONS} \vdash \mathrm{T}\ulcorner A \urcorner \to A$, we have the following.

FACT 9. *For any consistent fixed-point* $X = \mathscr{K}(X)$, $(\mathbb{N}, X) \vDash \mathrm{T}\ulcorner A \urcorner \to A$.

PROPOSITION 10. $(\mathbb{N}, \mathbb{T}_X) \vDash \mathsf{CD}^+_\mathrm{T}$.

Proof. By induction on the length of the proof in CD^+_T. We verify some key axioms, noting that T1 and T4–T6 are immediate by definition.

T2$^+$: $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \vee \mathrm{TF}\varphi$ iff $(\mathbb{N}, X) \vDash \mathrm{T}\varphi \vee \mathrm{F}\varphi$ iff $\{\varphi, \neg\varphi\} \cap X \neq \varnothing$ iff $\{\mathrm{T}\varphi, \mathrm{F}\varphi\} \cap X \neq \varnothing$ (by the fixed-point property), iff $(\mathbb{N}, X) \vDash \mathrm{TT}\varphi \vee \mathrm{TF}\varphi$ iff $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{T}(\mathrm{TT}\varphi \vee \mathrm{TF}\varphi)$.

T3: Assume $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \vee \mathrm{TF}\varphi$, which is the case iff $\{\varphi, \neg\varphi\} \cap X \neq \varnothing$. To show $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \to \mathrm{T}\varphi$, we reason as follows, letting $\varphi = \ulcorner A \urcorner$:[8]

$$\begin{aligned}
(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi &\quad \text{iff} \\
(\mathbb{N}, X) \vDash \mathrm{T}\varphi &\quad \textit{hence}, \text{ by Fact 9}, \\
(\mathbb{N}, X) \vDash A, &\quad \text{iff} \\
(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{T}\varphi. &
\end{aligned}$$

We notice that—due to the fact that we are reasoning in a consistent fixed-point model—the assumption $\mathrm{D}\varphi$ has not been employed in this part of the argument.

Conversely, to show that $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{T}\varphi \to \mathrm{TT}\varphi$, assume $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{T}\varphi$, which is equivalent to $(\mathbb{N}, X) \vDash A$. Towards a contradiction, suppose $(\mathbb{N}, \mathbb{T}_X) \nvDash \mathrm{TT}\varphi$, hence $(\mathbb{N}, X) \nvDash \mathrm{T}\varphi$. Then, since $\{\varphi, \neg\varphi\} \cap X \neq \varnothing$ by assumption, $(\mathbb{N}, X) \vDash \mathrm{F}\varphi$, and therefore $(\mathbb{N}, X) \vDash \neg A$ by Fact 9, which contradicts $(\mathbb{N}, X) \vDash A$.

[8] We assume that φ denotes a sentence without loss of generality, as $X \subseteq \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}$.

D3: We need to show that $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \vee \mathrm{TF}\varphi$ is equivalent to

$$(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\big(\mathrm{TT}\varphi \vee \mathrm{TF}\varphi\big) \vee \mathrm{TF}\big(\mathrm{TT}\varphi \vee \mathrm{TF}\varphi\big).$$

The latter is equivalent to

$$\underbrace{(\mathbb{N}, X) \vDash \mathrm{T}\big(\mathrm{TT}\varphi \vee \mathrm{TF}\varphi\big)}_{d_1} \text{ or } \underbrace{(\mathbb{N}, X) \vDash \mathrm{F}\big(\mathrm{TT}\varphi \vee \mathrm{TF}\varphi\big)}_{d_2}.$$

Since d_1 and d_2 are equivalent to, respectively, $(\mathbb{N}, X) \vDash \mathrm{T}\varphi \vee \mathrm{F}\varphi$ and $(\mathbb{N}, X) \vDash \mathrm{F}\varphi \wedge \mathrm{T}\varphi$, by consistency of X it can be observed that their disjunction is equivalent to d_1, which is equivalent to $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \vee \mathrm{TF}\varphi$, as required.

D5: Assume $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}(\varphi \wedge \psi) \vee \mathrm{TF}(\varphi \wedge \psi)$. If $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}(\varphi \wedge \psi)$, then $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \wedge \mathrm{TT}\psi$ by the closure properties of X, hence both φ and ψ are determinate. If $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TF}(\varphi \wedge \psi)$ then $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TF}\varphi \vee \mathrm{TF}\psi$ by the properties of X. If $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TF}\varphi$, then $(\mathbb{N}, X) \vDash \mathrm{F}\varphi$ and hence $(\mathbb{N}, X) \vDash \neg A$ by Fact 9, for $\varphi = \ulcorner A \urcorner$. This in turn yields $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TF}\varphi \wedge \mathrm{F}\varphi$, hence φ is determinate and false. Similarly if $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TF}\psi$.

Conversely, if $(\mathbb{N}, \mathbb{T}_X) \vDash (\mathrm{TT}\varphi \vee \mathrm{TF}\varphi) \wedge (\mathrm{TT}\psi \vee \mathrm{TF}\psi)$, then we get immediately $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}(\varphi \wedge \psi) \vee \mathrm{TF}(\varphi \wedge \psi)$ by closure properties of X. If $(\mathbb{N}, \mathbb{T}_X) \vDash (\mathrm{TT}\varphi \vee \mathrm{TF}\varphi) \wedge \mathrm{F}\varphi$, then $(\mathbb{N}, X) \vDash (\mathrm{T}\varphi \vee \mathrm{F}\varphi) \wedge \neg A$, which by consistency of X is equivalent to $(\mathbb{N}, X) \vDash \mathrm{F}\varphi$, hence $(\mathbb{N}, X) \vDash \mathrm{F}(\varphi \wedge \psi)$ for any ψ. It follows that $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TF}(\varphi \wedge \psi)$. Similarly if we assume $(\mathbb{N}, \mathbb{T}_X) \vDash (\mathrm{TT}\psi \vee \mathrm{TF}\psi) \wedge \mathrm{F}\psi$. □

It is clear that CD^+_T can interpret CD^+.

OBSERVATION 11. *There is an $\mathcal{L}_\mathbb{N}$-interpretation of CD^+ in CD^+_T.*

Proof. We employ the recursion theorem to define an $\mathcal{L}_\mathbb{N}$-translation $\delta\colon \mathcal{L}_\mathrm{D} \to \mathcal{L}_\mathrm{T}$ which systematically replaces D with its definition from (4). In more detail, the recursion theorem for primitive recursive functions[9] can be employed to define an $\mathcal{L}_\mathbb{N}$-translation $\delta\colon \mathcal{L}_\mathrm{D} \to \mathcal{L}_\mathrm{T}$ such that

$$(\mathrm{D}x)^\delta :\leftrightarrow \mathrm{T\underset{.}{T}}\mathrm{num}(\underset{.}{\delta}(x)) \vee \mathrm{T\underset{.}{F}}\mathrm{num}(\underset{.}{\delta}(x)).$$

The expression $\underset{.}{\delta}(x)$ abbreviates the formula representing δ in $\mathcal{L}_\mathbb{N}$. The verification that

$$\mathsf{CD}^+ \vdash A \text{ only if } \mathsf{CD}^+_\mathrm{T} \vdash A^\delta \tag{5}$$

is immediate given some basic syntactic facts, provable in a subtheory of PA, including:

$$\forall x(\mathrm{Sent}_{\mathcal{L}_\mathrm{D}}(x) \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(\underset{.}{\delta}(x))). \tag{6}$$

□

It will follow from the identity of $\mathsf{CKF_{cs}}$ and CD^+_T (Proposition 25) that CD^+ and CD^+_T are mutually $\mathcal{L}_\mathbb{N}$-interpretable. It would be too hasty to think, however, that CD^+ and CD^+_T are "notational variants"; CD^+ cannot define D in the manner prescribed by CD^+_T. This can be seen from the following observation. With reference to Fujimoto and Halbach's semantic construction for CD^+ described on page 7, let

$$T_B := T_\infty \cup \{B\},$$

9 See, for instance, [25, sec. 11.2].

for B an $\mathcal{L}_{\mathrm{D}}$-sentence, and

$$\mathbb{T}_B := \{A \in \mathcal{L}_{\mathrm{D}} \mid (\mathbb{N}, D_\infty, T_B) \vDash A\}.$$

LEMMA 12. *If* $B \notin D_\infty$, *then* $(\mathbb{N}, D_\infty, \mathbb{T}_B) \vDash \mathsf{CD}^+$.

Proof. We first show, by induction on A, that if $A \in D_\infty$, then

$$(\mathbb{N}, D_\infty, T_\infty) \vDash A \text{ iff } (\mathbb{N}, D_\infty, T_B) \vDash A. \tag{7}$$

The crucial case is when $A = \mathrm{T}\varphi$: $(\mathbb{N}, D_\infty, T_\infty) \vDash \mathrm{T}\varphi$ iff $\varphi \in T_\infty$, hence $(\mathbb{N}, D_\infty, T_B) \vDash \mathrm{T}\varphi$. For the converse direction, $(\mathbb{N}, D_\infty, T_B) \vDash \mathrm{T}\varphi$ iff either $\varphi \in T_\infty$, in which case we are done, or $\varphi = B$. However, since $\mathrm{T}\varphi \in D_\infty$, it follows that $\varphi \in D_\infty$, hence $\varphi \neq B$.

Having shown (7), we proceed with the main claim by induction on the length of proofs of CD^+. We check two cases where the use of (7) is relevant, letting $\varphi = \ulcorner A \urcorner$.

- For T3, assume $A \in D_\infty$. Then $(\mathbb{N}, D_\infty, T_\infty) \vDash A$ iff $A \in T_\infty$ since $(\mathbb{N}, D_\infty, \mathbb{T}_\infty)$ models CD^+. Then $(\mathbb{N}, D_\infty, \mathbb{T}_B) \vDash \mathrm{T}\varphi$ iff $(\mathbb{N}, D_\infty, T_B) \vDash A$ iff, by (7), $(\mathbb{N}, D_\infty, T_\infty) \vDash A$ iff $A \in T_\infty$ iff, since $A \neq B$, $(\mathbb{N}, D_\infty, T_B) \vDash \mathrm{T}\varphi$ iff $(\mathbb{N}, D_\infty, \mathbb{T}_B) \vDash \mathrm{TT}\varphi$.
- For D5, the left-to-right direction is immediate. As for right-to-left direction, it suffices to consider the case where, e.g., $(\mathbb{N}, D_\infty, \mathbb{T}_B) \vDash \mathrm{D}\varphi \wedge \mathrm{F}\varphi$. This is the case iff $(\mathbb{N}, D_\infty, T_B) \vDash \mathrm{D}\varphi \wedge \neg A$ iff, by (7), $(\mathbb{N}, D_\infty, T_\infty) \vDash \mathrm{D}\varphi \wedge \neg A$ iff $(\mathbb{N}, D_\infty, \mathbb{T}_\infty) \vDash \mathrm{D}\varphi \wedge \mathrm{F}\varphi$, hence $\varphi \wedge \psi \in D_\infty$ for any ψ, hence $(\mathbb{N}, D_\infty, \mathbb{T}_B) \vDash \mathrm{D}(\varphi \wedge \psi)$. □

While $\mathsf{CD}^+ \vdash (\forall \varphi \colon \mathcal{L}_{\mathrm{D}})(\mathrm{D}\varphi \to \mathrm{TT}\varphi \vee \mathrm{TF}\varphi)$, Lemma 12 yields that the converse implication is not provable.

COROLLARY 13. $\mathsf{CD}^+ \nvdash (\forall \varphi \colon \mathcal{L}_{\mathrm{D}})(\mathrm{TT}\varphi \vee \mathrm{TF}\varphi \to \mathrm{D}\varphi)$.

COROLLARY 14. CD^+ does not *prove any of the following sentences*:

(i) $\forall t(\mathrm{TTT}t \to \mathrm{TT}t)$;
(ii) $\forall t(\mathrm{TT}t \to \mathrm{TTT}t)$;
(iii) $(\forall \varphi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TT}\varphi \leftrightarrow \mathrm{TT}\neg\neg\varphi)$;
(iv) $(\forall \varphi, \psi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TT}(\varphi \wedge \psi) \leftrightarrow \mathrm{TT}\varphi \wedge \mathrm{TT}\psi)$;
(v) $(\forall \varphi, \psi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TF}(\varphi \wedge \psi) \leftrightarrow \mathrm{TF}\varphi \vee \mathrm{TT}\psi)$;
(vi) $(\forall \varphi(v) \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TT}\forall\varphi(v) \leftrightarrow \mathrm{T}\forall t \mathrm{T}\varphi(t)$;
(vii) $(\forall \varphi(v) \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TF}\forall\varphi(v) \leftrightarrow \mathrm{T}\exists t \mathrm{F}\varphi(t)$.

Proof. For (i), to construct the required countermodel use $\neg\lambda$, for λ a liar sentence, in place of B in Lemma 12; the model invalidates $\mathrm{TTT}\lambda \to \mathrm{TT}\lambda$. For (ii), use λ in place of B so to invalidate $\mathrm{TT}\lambda \to \mathrm{TTT}\lambda$. The other cases can be dealt with by emploing suitable truth-teller sentences. For instance, for (iv), define parametrized truth-tellers τ_0 and τ_1, and replace $\tau_0 \wedge \tau_1$ for B in Lemma 12. □

§4. KF's classical closure. The definition of $\mathsf{CD}^+_{\mathrm{T}}$ is parasitic on Fujimoto and Halbach's CD^+. As such, they are axiomatizations of *both truth and determinateness*. It is then natural to ask whether there is an *axiomatization of truth* that can be directly inspired by the models $(\mathbb{N}, \mathbb{T}_X)$ introduced above and yet deliver the required principles for determinateness. In such models, full disquotation is allowed only under two or more layers of truth (or falsity): outside those layers, fully classical principles are licensed at the expense of disquotation. We will now answer the question positively. We will then explain in the final section how the theories capture the conception of truth (and determinateness) outlined in the opening section.

We introduce the theory $\mathsf{CKF_{cs}}$ (standing for the classical closure of $\mathsf{KF} + \mathsf{CONS}$). $\mathsf{CKF_{cs}}$ combines full compositionality (T1 and T4–T6) with Kripkean truth conditions in the inner layers of truth, along with a disquotation principle for truth ascriptions that will ensure the consistency of the inner truth predicate.

Definition 15 ($\mathsf{CKF_{cs}}$). *The axioms of the $\mathcal{L}_{\mathrm{T}}$-theory* $\mathsf{CKF_{cs}}$ *are the axioms of* PAT *together with*:

$$\forall s \forall t (\mathrm{T}(s = t) \leftrightarrow s^\circ = t^\circ) \tag{T1}$$
$$(\forall \varphi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{T}(\neg\varphi) \leftrightarrow \neg \mathrm{T}\varphi) \tag{T4}$$
$$(\forall \varphi, \psi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{T}(\varphi \wedge \psi) \leftrightarrow \mathrm{T}\varphi \wedge \mathrm{T}\psi) \tag{T5}$$
$$(\forall \varphi(v) \colon \mathcal{L}_{\mathrm{T}})(\mathrm{T}(\forall v \varphi) \leftrightarrow \forall t\, \mathrm{T}\varphi(t/v)) \tag{T6}$$
$$\forall s \forall t (\mathrm{TT}s = t \leftrightarrow s^\circ = t^\circ) \tag{T7}$$
$$\forall s \forall t (\mathrm{TF}s = t \leftrightarrow s^\circ \neq t^\circ) \tag{T8}$$
$$(\forall \varphi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TT}\varphi \leftrightarrow \mathrm{TT}\neg\neg\varphi) \tag{T9}$$
$$(\forall \varphi, \psi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TT}(\varphi \wedge \psi) \leftrightarrow \mathrm{TT}\varphi \wedge \mathrm{TT}\psi) \tag{T10}$$
$$(\forall \varphi, \psi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TF}(\varphi \wedge \psi) \leftrightarrow \mathrm{TF}\varphi \vee \mathrm{TT}\psi) \tag{T11}$$
$$(\forall \varphi(v) \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TT}\forall \varphi(v) \leftrightarrow \mathrm{T}\forall t \mathrm{T}\varphi(t)) \tag{T12}$$
$$(\forall \varphi(v) \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TF}\forall \varphi(v) \leftrightarrow \mathrm{T}\exists t \mathrm{F}\varphi(t)) \tag{T13}$$
$$\forall t (\mathrm{TTT}t \leftrightarrow \mathrm{TT}t) \tag{T14}$$
$$\forall t (\mathrm{TFT}t \leftrightarrow \mathrm{TF}t) \tag{T15}$$
$$\forall t (\mathrm{TT}t \rightarrow \mathrm{T}t^\circ) \tag{TDel}$$
$$(\forall \varphi(v) \colon \mathcal{L}_{\mathrm{T}}) \forall s \forall t \big(s^\circ = t^\circ \rightarrow (\mathrm{T}\varphi(s) \leftrightarrow \mathrm{T}\varphi(t))\big). \tag{R1}$$

Remark 16. *In the definition of* $\mathsf{CKF_{cs}}$, *the version of* KF *that 'lives' inside one layer of truth is slightly different from the one presented in Definition 3. Specifically, axiom* T15 *of* $\mathsf{CKF_{cs}}$ *does not feature, internally, the extra disjunct of* $\mathsf{KF2}$. *This reformulation, albeit inessential either conceptually or from the perspective of proof-theoretic strength, is required by our arguments to obtain the identity of* $\mathsf{CKF_{cs}}$ *and* $\mathsf{CD}^+_{\mathrm{T}}$ *established in Proposition 25.*

We list some basic theorems of $\mathsf{CKF_{cs}}$ that will be explicitly used below.[10]

Observation 17. *The following are derivable in* $\mathsf{CKF_{cs}}$:

$$\forall t (\mathrm{TF}t \leftrightarrow \mathrm{TTF}t) \tag{8}$$
$$\mathrm{TF}\varphi \rightarrow \neg \mathrm{TT}\varphi \tag{TCons}$$
$$(\forall \varphi, \psi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TT}(\varphi \vee \psi) \leftrightarrow \mathrm{TT}\varphi \vee \mathrm{TT}\psi) \tag{9}$$
$$(\forall \varphi, \psi \colon \mathcal{L}_{\mathrm{T}})(\mathrm{TF}(\varphi \vee \psi) \leftrightarrow \mathrm{TF}\varphi \wedge \mathrm{TF}\psi). \tag{10}$$

[10] The statement TCons amounts to the consistency axiom Cons under one additional layer of truth. Similarly to TDel, it ensures the consistency of the inner truth predicate. We will come back to this below, §5.

As mentioned, the theory $\mathsf{CKF_{cs}}$ is directly inspired by the models $(\mathbb{N}, \mathbb{T}_X)$ introduced in the previous section. Our next results provide a full characterization of standard models of $\mathsf{CKF_{cs}}$ in these terms: we show that they are exactly the classical closures of consistent fixed points.

Lemma 18. *For any consistent fixed-point X, $(\mathbb{N}, \mathbb{T}_X) \vDash \mathsf{CKF_{cs}}$.*

Proof. Axioms T1 and T4–T6 are immediate by definition. Axioms T7–T15 and R1 readily follow from properties of X. As for T16, we reason as in the argument for case T3 in the proof of Proposition 10. □

For any set S, let $S^{\mathscr{I}} := \{A \mid (\mathbb{N}, S) \vDash \mathrm{T}\ulcorner\mathrm{T}\ulcorner A\urcorner\urcorner\}$. Then

Lemma 19. *Let $(\mathbb{N}, S) \vDash \mathsf{CKF_{cs}}$. Then $S^{\mathscr{I}}$ is a consistent fixed point.*

Proof. Let $(\mathbb{N}, S) \vDash \mathsf{CKF_{cs}}$. By induction on the positive complexity of A, one shows that $A \in S^{\mathscr{I}}$ iff $A \in \mathscr{K}(S^{\mathscr{I}})$. For example, if A is of the form $\neg\mathrm{T}t$, one uses T15, or if A is of the form $\forall x B$, one uses T6 and T12. □

Lemma 20. *Let $(\mathbb{N}, S) \vDash \mathsf{CKF_{cs}}$. Then $S = \mathbb{T}_{S^{\mathscr{I}}}$.*

Proof. Let $(\mathbb{N}, S) \vDash \mathsf{CKF_{cs}}$. Again by positive induction on A, one shows that $(\mathbb{N}, S) \vDash \mathrm{T}\ulcorner A\urcorner$ iff $(\mathbb{N}, S^{\mathscr{I}}) \vDash A$. For example, if A is $\mathrm{T}\ulcorner B\urcorner$, then $(\mathbb{N}, S) \vDash \mathrm{T}\ulcorner\mathrm{T}\ulcorner B\urcorner\urcorner$ iff $B \in S^{\mathscr{I}}$ iff $(\mathbb{N}, S^{\mathscr{I}}) \vDash A$. □

From Lemmata 18–20, we obtain the following characterization theorem.

Proposition 21 ($\mathbb{N}$-Categoricity). $(\mathbb{N}, S) \vDash \mathsf{CKF_{cs}}$ *iff* $S = \mathbb{T}_{S^{\mathscr{I}}}$.

There's also a precise sense in which the logical strength of $\mathsf{CKF_{cs}}$ coincides with the one of $\mathsf{CT[KF + CONS]}$, and therefore of CD^+.

Proposition 22. $\mathsf{CKF_{cs}}$, $\mathsf{CT[KF + CONS]}$, *and* CD^+ *are mutually $\mathcal{L}_{\mathbb{N}}$-interpretable.*

Proof. To $\mathcal{L}_{\mathbb{N}}$-interpret $\mathsf{CKF_{cs}}$ in $\mathsf{CT[KF + CONS]}$, one defines an $\mathcal{L}_{\mathbb{N}}$-translation $\tau\colon \mathcal{L}_{\mathrm{T}} \to \mathcal{L}_{\mathrm{T},\mathbf{T}}$ that replaces outer occurrences of T with $\mathbf{T}$:

$$\begin{aligned}
\tau(s = t) &:\leftrightarrow s = t\\
\tau(\mathrm{T}x) &:\leftrightarrow \mathbf{T}x\\
\tau(\neg A) &:\leftrightarrow \neg\tau(A)\\
\tau(A \wedge B) &:\leftrightarrow \tau(A) \wedge \tau(B)\\
\tau(\forall v A) &:\leftrightarrow \forall x\, \tau(A(x/v)).
\end{aligned}$$

The translation τ is only external, and does not require applications of the primitive recursion theorem in its definition. The verification that τ is an $\mathcal{L}_{\mathbb{N}}$-interpretation is fairly straightforward. The compositional principles (T1 and T4–T6) follow from the definition of the translation τ and the compositional axioms for $\mathbf{T}$ of $\mathsf{CT[KF + CONS]}$. Axioms T7–T13 follow from the definition of τ, $\mathbf{T}1$, and the corresponding axiom of KF. For instance, for T7 (in non-abbreviated form):

$$\begin{aligned}
\tau(\mathrm{T}\underset{\cdot}{\mathrm{T}}\mathrm{num}(s \underset{\cdot}{=} t) \leftrightarrow s^\circ = t^\circ) &\text{ iff } \mathbf{T}\underset{\cdot}{\mathrm{T}}\mathrm{num}(s \underset{\cdot}{=} t) \leftrightarrow s^\circ = t^\circ && \text{def. of } \tau\\
&\text{ iff } \mathrm{T}(s \underset{\cdot}{=} t) \leftrightarrow s^\circ = t^\circ && \text{by } \mathbf{T}1.
\end{aligned}$$

The last line is an axiom of $\mathsf{CT[KF + CONS]}$. Axiom R1 is proved by formal induction on the complexity of $\varphi(v)$.

For TDel, one notices that, by the definition of τ and $\mathbf{T}1$, the translation of TDel becomes

$$\mathrm{T}t^\circ \to \mathbf{T}t^\circ. \tag{11}$$

Now either $\neg\mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$ or $\mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$. In the latter case we can then prove (11) by formal induction on the positive complexity of t°. Notice that in the case in which t° is of form $\neg\mathrm{T}s$, we employ Cons. If $\neg\mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$, then $\mathrm{FT}t$ by KF2, and $\neg\mathrm{TT}t$ by Cons, thus $\neg\mathrm{T}t^\circ$ and $\mathrm{T}t^\circ \to \mathbf{T}t^\circ$ follows trivially.

To interpret CT[KF + CONS] in $\mathsf{CKF}_{\mathsf{cs}}$, we first define the internal translation $\imath\colon \mathcal{L}_{\mathrm{T},\mathbf{T}} \to \mathcal{L}_\mathrm{T}$ operating on codes of $\mathcal{L}_{\mathrm{T},\mathbf{T}}$-formulae as follows:

$$\begin{aligned}
\imath(\varphi) &:= \varphi, && \text{for } A \text{ atomic of } \mathcal{L}_\mathbb{N}\\
\imath(\psi) &:= \ulcorner 0 = 1 \urcorner, && \text{for } \psi \in \mathrm{Sent}_{\mathcal{L}_{\mathrm{T},\mathbf{T}}} \setminus \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}\\
\imath(\mathrm{T}x) &:= \mathrm{T}\underset{\cdot}{\imath}(x)\\
\imath(\neg\varphi) &:= \neg\imath(\varphi)\\
\imath(\varphi \wedge \psi) &:= \imath(\varphi) \wedge \imath(\psi)\\
\imath(\forall v\varphi) &:= \forall x\imath(\varphi(x/v)).
\end{aligned}$$

We then define the full translation $\sigma\colon \mathcal{L}_{\mathrm{T},\mathbf{T}} \to \mathcal{L}_\mathrm{T}$, which replaces $\mathbf{T}$ with one single layer of $\mathsf{CKF}_{\mathsf{cs}}$-truth and T by two, and behaves internally according to $\imath$:

$$\begin{aligned}
\sigma(A) &:\leftrightarrow A, && \text{for } A \text{ atomic of } \mathcal{L}_\mathbb{N}\\
\sigma(\mathbf{T}x) &:\leftrightarrow \mathrm{T}\underset{\cdot}{\imath}(x)\\
\sigma(\mathrm{T}x) &:\leftrightarrow \mathrm{T}\underset{\cdot}{\mathrm{T}}\mathrm{num}(\underset{\cdot}{\imath}(x))\\
\sigma(\neg A) &:\leftrightarrow \neg\sigma(A)\\
\sigma(A \wedge B) &:\leftrightarrow \sigma(A) \wedge \sigma(B)\\
\sigma(\forall vA) &:\leftrightarrow \forall x\sigma(A(x/v)).
\end{aligned}$$

It remains to verify that σ is the required $\mathcal{L}_\mathbb{N}$-interpretation of CT[KF + CONS] in $\mathsf{CKF}_{\mathsf{cs}}$. We consider the crucial case of KF2. The translation of its first conjunct is

$$\mathrm{T}\underset{\cdot}{\mathrm{T}}\mathrm{num}(\underset{\cdot}{\mathrm{T}}\underset{\cdot}{\imath}(t)) \leftrightarrow \mathrm{T}\underset{\cdot}{\mathrm{T}}(\underset{\cdot}{\imath}(t)), \tag{12}$$

which follows immediately from T14. The translation of the second conjunct of KF2 is

$$\mathrm{T}\underset{\cdot}{\mathrm{T}}\mathrm{num}\underset{\cdot}{\neg}\underset{\cdot}{\mathrm{T}}\underset{\cdot}{\imath}(t) \leftrightarrow \mathrm{T}\underset{\cdot}{\mathrm{T}}(\underset{\cdot}{\neg}\underset{\cdot}{\imath}(t)) \vee \neg\mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ). \tag{13}$$

The left-to-right direction of (13) follows directly from T15. Similarly for the right-to-left direction, if $\mathrm{T}\underset{\cdot}{\mathrm{T}}(\underset{\cdot}{\neg}\underset{\cdot}{\imath}(t))$ holds. If $\neg\mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$, $\imath(t^\circ) = \ulcorner 0 = 1 \urcorner$, so $\mathrm{T}\underset{\cdot}{\mathrm{T}}\mathrm{num}(\underset{\cdot}{\mathrm{T}}\underset{\cdot}{\neg}\underset{\cdot}{\imath}(t))$. □

Next, we turn to the question how CD^+_T and $\mathsf{CKF}_{\mathsf{cs}}$ are related. Corollaries 13 and 14 show that some of the axioms of $\mathsf{CKF}_{\mathsf{cs}}$ are not provable in CD^+. However, as we shall see shortly, *these are all provable* in CD^+_T. In fact, we shall see that CD^+_T and $\mathsf{CKF}_{\mathsf{cs}}$ are identical theories.

LEMMA 23. CD^+_T *is a subtheory of* $\mathsf{CKF}_{\mathsf{cs}}$.

Proof. We verify a few key axioms, reasoning informally within $\mathsf{CKF}_{\mathsf{cs}}$.
For $\mathrm{T}2^+$, we use T14, (8), and distribution of T over $\vee$.

For T3, since $\mathrm{TT}t \to \mathrm{T}t^\circ$ is an axiom of $\mathsf{CKF_{cs}}$, we have $\mathrm{TT}t \to (\mathrm{TT}t \leftrightarrow \mathrm{T}t^\circ)$. So assume $\mathrm{TF}t$. Then $\mathrm{F}t^\circ$, hence $\neg\mathrm{T}t^\circ$ and therefore $\mathrm{T}t^\circ \to \mathrm{TT}t$, hence $\mathrm{T}t^\circ \leftrightarrow \mathrm{TT}t$.

For D3, we use TCons and (10)—derivable in $\mathsf{CKF_{cs}}$—along with axioms T14 and T15.

For D5, assume $\mathrm{TT}(\varphi \wedge \psi) \vee \mathrm{TF}(\varphi \wedge \psi)$. By T10 and T11, this is equivalent to

$$(\mathrm{TT}\varphi \wedge \mathrm{TT}\psi) \vee (\mathrm{TF}\varphi \vee \mathrm{TF}\psi).$$

Since we have $\mathrm{TF}t \to \mathrm{F}t^\circ$, by a series of propositional inferences we obtain

$$\underbrace{((\mathrm{TT}\varphi \vee \mathrm{TF}\varphi) \wedge (\mathrm{TT}\psi \vee \mathrm{TF}\psi))}_{\delta_1} \vee \underbrace{((\mathrm{TT}\varphi \vee \mathrm{TF}\varphi) \wedge \mathrm{F}\varphi)}_{\delta_2} \vee \underbrace{((\mathrm{TT}\psi \vee \mathrm{TF}\psi) \wedge \mathrm{F}\psi)}_{\delta_3}.$$

Conversely, each combination obtainable from the conjuncts of δ_1 entails $\mathrm{TT}(\varphi \wedge \psi) \vee \mathrm{TF}(\varphi \wedge \psi)$. As δ_2 and δ_3, use T16 and T4 to obtain $\mathrm{TF}\varphi \wedge \mathrm{F}\varphi$, respectively, $\mathrm{TF}\psi \wedge \mathrm{F}\psi$, which entail the desired conclusion via T11. □

LEMMA 24. $\mathsf{CKF_{cs}}$ *is a subtheory of* $\mathsf{CD}^+_{\mathrm{T}}$.

Proof. The key observation is that, since $\mathsf{CD}^+_{\mathrm{T}} \vdash \mathrm{TT}\varphi \vee \mathrm{TF}\varphi \to (\mathrm{T}\varphi \leftrightarrow \mathrm{TT}\varphi)$, we can perform the necessary quotation and disquotation steps to prove compositionality within layers of T. For example, to see that $\mathsf{CD}^+_{\mathrm{T}} \vdash$ T9, assume $\mathrm{TT}\varphi$ (or $\mathrm{TT}\neg\neg\varphi$). Then $\mathrm{TT}\varphi \vee \mathrm{TF}\varphi$ and $\mathrm{TT}\neg\varphi \vee \mathrm{TF}\neg\varphi$, therefore we have both $\mathrm{TT}\varphi \leftrightarrow \mathrm{T}\varphi$ and $\mathrm{TT}\neg\neg\varphi \leftrightarrow \mathrm{T}\neg\neg\varphi$. But then $\mathrm{TT}\varphi \leftrightarrow \mathrm{T}\varphi \leftrightarrow \mathrm{T}\neg\neg\varphi \leftrightarrow \mathrm{TT}\neg\neg\varphi$. The derivability of other axioms follows a similar pattern. We show some examples, reasoning informally within $\mathsf{CD}^+_{\mathrm{T}}$.

T10: Assume $\mathrm{TT}(\varphi \wedge \psi)$. Then $\mathrm{TT}(\varphi \wedge \psi) \to \mathrm{T}(\varphi \wedge \psi)$ by T3,[11] hence $\mathrm{T}\varphi \wedge \mathrm{T}\psi$, and therefore $\neg\mathrm{F}\varphi \wedge \neg\mathrm{F}\psi$. This together with $\mathrm{TT}(\varphi \wedge \psi)$ yields, via D5, that φ and ψ are determinate, i.e., $\mathrm{TT}\varphi \vee \mathrm{TF}\varphi$ and $\mathrm{TT}\psi \vee \mathrm{TF}\psi$. Hence we conclude $\mathrm{TT}\varphi \wedge \mathrm{TT}\psi$ from $\mathrm{T}\varphi \wedge \mathrm{T}\psi$.

For the converse direction, $\mathrm{TT}\varphi \wedge \mathrm{TT}\psi$ entails $\mathrm{TT}(\varphi \wedge \psi) \vee \mathrm{TF}(\varphi \wedge \psi)$ by D5. Moreover, since $\mathrm{TT}\varphi \wedge \mathrm{TT}\psi$ also entails $\mathrm{T}\varphi \wedge \mathrm{T}\psi$, we get $\mathrm{T}(\varphi \wedge \psi)$, hence $\mathrm{TT}(\varphi \wedge \psi)$ by T3.

T16: If $\mathrm{TT}\varphi$, then $\mathrm{TT}\varphi \vee \mathrm{TF}\varphi$, hence $\mathrm{T}\varphi$ by T3. □

As a corollary we obtain the following.

PROPOSITION 25. $\mathsf{CD}^+_{\mathrm{T}}$ *and* $\mathsf{CKF_{cs}}$ *are identical theories.*

§5. Alternative axiomatizations.

As mentioned, the theory $\mathsf{CKF_{cs}}$ contains, along with Kripkean truth conditions in the inner layers of truth, an axiom (TDel) restricting the class of fixed points to those that are consistent. In this section, we discuss alternative consistency axioms.

The axiom TDel is reminiscent of the schema often called T-Out: $\mathrm{T}\ulcorner A \urcorner \to A$. It is also known (see [4]) that T-Out is equivalent to the consistency axiom Cons: $\forall\varphi : \mathcal{L}_{\mathrm{T}}(\mathrm{F}\varphi \to \neg\mathrm{T}\varphi)$. It may be asked whether TDel is equivalent to TCons from Observation 17: $\mathrm{TF}\varphi \to \neg\mathrm{TT}\varphi$. The fact that the latter follows from the former is straightforward: $\mathrm{TF}\varphi \to \mathrm{F}\varphi \to \neg\mathrm{T}\varphi \to \neg\mathrm{TT}\varphi$. However, the converse employs an extra axiom stating that only sentences are truly true.

LEMMA 26. $\mathsf{CKF_{cs}} - \mathrm{TDel} + \mathrm{TCons} + \mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_{\mathrm{T}}}(t^\circ) \vdash \mathrm{TDel}$.

[11] Under the definition $\mathrm{D}x :\leftrightarrow \mathrm{TT}x \vee \mathrm{TF}x$. Same remark applies below.

Proof. We distinguish cases. If $t^\circ \notin \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}$, then $\neg\mathrm{TT}t$, hence trivially $\mathrm{TT}t \to \mathrm{T}t$. Else, we reason by induction on the formal complexity of $t^\circ = \varphi$. If φ is an equality, the claim follows from T7 and T8. If $\varphi \equiv \mathrm{T}t$, the claim follows from T14. If $\varphi \equiv \neg\mathrm{T}t$, we use T15 and the additional axiom: $\mathrm{TT}\neg\mathrm{T}t \to \mathrm{TF}t \to \neg\mathrm{TT}t \to \mathrm{T}\neg\mathrm{T}t$. $\square$

Additionally, it can be observed that TCons is equivalent to $\mathrm{TFT}t \to \mathrm{FT}t$.

Lemma 27. *The theories* $\mathsf{CKF_{cs}} - \mathrm{TDel} + \mathrm{TCons}$ *and* $\mathsf{CKF_{cs}} - \mathrm{TDel} + \mathrm{TF}t \to \neg\mathrm{TT}t$ *prove the same theorems.*

Proof. To derive TCons from $\mathrm{TFT}t \to \mathrm{FT}t$, we use T15: $\mathrm{TF}t \to \mathrm{TFT}t \to \mathrm{FT}t \to \neg\mathrm{TT}t$. Conversely, the claim follows from the previous lemma observing that, for $t \in \mathrm{CTerm}$, we have $\mathrm{TF}t \in \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}$. $\square$

Collecting these observation together, we obtain the following proposition (cf. [16, lemma 15.9]).

Proposition 28. *Over* $\mathsf{CKF_{cs}} - \mathrm{TDel} + \mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$, *the following statements are equivalent*:

(i) $\mathrm{TT}t \to \mathrm{T}t$;
(ii) $\mathrm{TF}t \to \neg\mathrm{TT}t$;
(iii) $\mathrm{TFT}t \to \mathrm{FT}t$.

Some formulations of KF (e.g., [3, 23]) do include the axiom $\mathrm{T}t \to \mathrm{Sent}(t^\circ)$. However, we decided not to include $\mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$ in our official formulation of $\mathsf{CKF_{cs}}$ in order to simplify its comparison with CD^+_T. If $\mathsf{CKF_{cs}}$ were defined without TDel but with, for example, TCons along with $\mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$, then the equivalence stated in Proposition 25 would need to be reformulated as follows.

Lemma 29. *The theories* $\mathsf{CKF}'_{\mathsf{cs}} := \mathsf{CKF_{cs}} - \mathrm{TDel} + \mathrm{TCons} + \mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$ *and* $\mathsf{CD}^+_\mathrm{T} + \mathrm{D}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$ *are identical.*[12]

Proof. For the inclusion of $\mathsf{CD}^+_\mathrm{T} + \mathrm{D}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t)$ into $\mathsf{CKF}'_{\mathsf{cs}}$, the crucial axioms are T3 and $\mathrm{D}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$. As for the latter, $\mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$ is just an axiom of $\mathsf{CKF}'_{\mathsf{cs}}$. As for $\mathrm{TF}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$, it follows from $\mathrm{TT}\dot\neg t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(\dot\neg t^\circ)$, hence $\mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$.

For T3, we distinguish two cases. If $t^\circ \notin \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}$, then $\dot\neg t^\circ \notin \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}$, hence $\mathrm{TT}t \to \bot \wedge \mathrm{TF}t \to \bot$, hence trivially $\mathrm{TT}t \vee \mathrm{TF}t \to (\mathrm{TT}t \leftrightarrow \mathrm{T}t^\circ)$. If $t \in \mathrm{Sent}(t)$, by Lemma 26 we have $\mathrm{TT}t \to \mathrm{T}t^\circ$, hence $\mathrm{TT}t \leftrightarrow (\mathrm{T}t^\circ \leftrightarrow \mathrm{TT}t)$. For the second disjunct, using Proposition 28 we have $\mathrm{TF}t \to \mathrm{F}t^\circ \to \neg\mathrm{T}t^\circ \to (\mathrm{T}t^\circ \to \mathrm{TT}t)$.

For the converse inclusion of $\mathsf{CKF}'_{\mathsf{cs}}$ into $\mathsf{CD}^+_\mathrm{T} + \mathrm{D}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t)$, the crucial cases are TCons and $\mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$. The latter follows from $\mathrm{D}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t)$. The former can be derived thus: $\mathrm{TF}t \to \mathrm{F}t \to \neg\mathrm{T}t \to \neg\mathrm{TT}t$. $\square$

§6. Complete, symmetric, and mixed fixed points. In this section, we verify whether the results from previous sections carry over if one focuses on different classes of fixed points of $\mathscr{K}$. We provide a positive answer for the class of complete fixed points as well as for the class of consistent *or* complete fixed points. Specifically, given a complete fixed-point X of $\mathscr{K}$, the structure $(\mathbb{N}, \mathbb{T}_X)$ can be shown to be a model of the

[12] The same would hold for $\mathsf{CD}^+_\mathrm{T} + \mathrm{TT}t \to \mathrm{Sent}_{\mathcal{L}_\mathrm{T}}(t^\circ)$.

 https://doi.org/10.1017/S1755020325100968

determinateness axioms of CD^+ with D defined in terms of T. Moreover, the resulting theory can be shown to be identical to a variant of $\mathsf{CKF}_{\mathsf{cs}}$. Similarly for the class of fixed points which are either consistent or complete. The question whether similar results are available for mixed fixed points, i.e., fixed points which are neither consistent nor complete, will be left open.

6.1. Complete models. Since the structure of the arguments is very similar, we limit ourselves to highlighting the necessary modifications.

DEFINITION 30 ($\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$). *Let* $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$ *be the* $\mathcal{L}_{\mathrm{T}}$*-theory whose axioms are those of* CD^+*, but where D defined in terms of* T *as*

$$\mathrm{D}x :\leftrightarrow \neg\mathrm{TT}x \vee \neg\mathrm{TF}x. \tag{14}$$

It can be shown that, given a complete fixed-point $X = \mathscr{K}(X)$, the structure $(\mathbb{N}, \mathbb{T}_X)$, where $\mathbb{T}_X := \{A \in \mathcal{L}_{\mathrm{T}} \mid (\mathbb{N}, X) \vDash A\}$, is a model of $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$. The argument follows the blueprint of Proposition 10. We first observe that, since a complete fixed-point X of $\mathscr{K}$ is such that $(\mathbb{N}, X) \vDash \mathsf{KF} + \mathsf{COMP}$, and since $\mathsf{KF} + \mathsf{COMP}$ derives the schema $A \to \mathrm{T}\ulcorner A \urcorner$, we have the following.

FACT 31. *For any complete fixed-point X, for any A, $(\mathbb{N}, X) \vDash A \to \mathrm{T}\ulcorner A \urcorner$.*

Just as Fact 9 was used in the proof of Proposition 10, Fact 31 will play a similar role in the proof of the following.

PROPOSITION 32. *For a complete fixed-point X, $(\mathbb{N}, \mathbb{T}_X) \vDash \mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$.*

Proof Sketch. By induction on the length of proofs in $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$. We verify T3 and D3, whose arguments are symmetric to those in the proof of Proposition 10.

T3: Assume $(\mathbb{N}, \mathbb{T}_X) \vDash \neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi$, which is the case iff $\{\varphi, \neg\varphi\} \not\subseteq X$. To show $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \to \mathrm{T}\varphi$, let $\varphi = \ulcorner A \urcorner$ and assume $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi$, which is equivalent to $\varphi \in X$. Towards a contradiction, suppose $(\mathbb{N}, \mathbb{T}_X) \nvDash \mathrm{T}\varphi$, hence $(\mathbb{N}, X) \nvDash A$. By Fact 31, we get $(\mathbb{N}, X) \vDash \mathrm{F}\varphi$ iff $\neg\varphi \in X$, contradicting our assumption.

Conversely, to show $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{T}\varphi \to \mathrm{TT}\varphi$, we reason as follows without using the assumption on the determinateness of φ:

$$\begin{array}{ll} (\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{T}\varphi, & \text{iff} \\ (\mathbb{N}, X) \vDash A, & \textit{hence}, \text{ by Fact 31} \\ (\mathbb{N}, X) \vDash \mathrm{T}\varphi, & \text{iff} \\ (\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi. & \end{array}$$

D3: We need to show that $(\mathbb{N}, \mathbb{T}_X) \vDash \neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi$ is equivalent to

$$(\mathbb{N}, \mathbb{T}_X) \vDash \neg\mathrm{TT}(\neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi) \vee \neg\mathrm{TF}(\neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi).$$

The latter is equivalent to

$$\underbrace{(\mathbb{N}, X) \vDash \neg\mathrm{T}(\neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi)}_{d_1} \quad \text{or} \quad \underbrace{(\mathbb{N}, X) \vDash \neg\mathrm{F}(\neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi)}_{d_2}.$$

Since d_1 and d_2 are equivalent to, respectively, $(\mathbb{N}, X) \vDash \neg\mathrm{T}\varphi \wedge \neg\mathrm{F}\varphi$ and $(\mathbb{N}, X) \vDash \neg\mathrm{F}\varphi \vee \neg\mathrm{T}\varphi$, by completeness of X it can be observed that their disjunction is equivalent to d_2, which is equivalent to $(\mathbb{N}, \mathbb{T}_X) \vDash \neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi$, as required. □

A collection of principles inspired by the models $(\mathbb{N}, \mathbb{T}_X)$ for X a complete fixed point can be obtained by modifying $\mathsf{CKF_{cs}}$ in the obvious way, that is, by replacing the axiom expressing consistency with one expressing completeness, leaving the remaining axioms characterising the closure of KF unmodified.

Definition 33 ($\mathsf{CKF_{cp}}$). $\mathsf{CKF_{cp}}$ *is the system obtained from* $\mathsf{CKF_{cs}}$ *by replacing TDel with*

$$\mathrm{T}t \rightarrow \mathrm{TT}t. \tag{TRep}$$

Axiom TRep readily yields completeness $\mathrm{TT}t \vee \mathrm{TF}t$ via $\neg\mathrm{TT}t \rightarrow \neg\mathrm{T}t^\circ \rightarrow \mathrm{F}t^\circ \rightarrow \mathrm{TF}t$.

Proposition 34. $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$ *and* $\mathsf{CKF_{cp}}$ *are identical theories.*

Proof Sketch. As in the proof of Lemma 24, the key observation to show that $\mathsf{CKF_{cp}}$ is a subtheory of $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$ is that $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}] \vdash \neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi \rightarrow (\mathrm{T}\varphi \leftrightarrow \mathrm{TT}\varphi)$. We can derive the counterpositive of each axiom of $\mathsf{CKF_{cp}}$ by performing the necessary quotation and disquotation step. For example, for T10 we assume $\neg\mathrm{TT}(\varphi \wedge \psi)$, which implies $\neg\mathrm{T}(\varphi \wedge \psi)$, hence $\neg\mathrm{T}\varphi \vee \neg\mathrm{T}\psi$. Via D5, at least one between φ and ψ is determinate, hence we conclude $\neg\mathrm{TT}\varphi \vee \neg\mathrm{TT}\psi$. In a similar way we can derive the counterpositive of TRep: if $\neg\mathrm{TT}t$, then $\mathrm{T}t \leftrightarrow \mathrm{TT}t$, hence $\neg\mathrm{T}t$.

Conversely, to show that $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$ is a subtheory of $\mathsf{CKF_{cp}}$, the reasoning is similar to the proof for Lemma 23. We only verify T3: Since $\mathrm{T}t \rightarrow \mathrm{TT}t$, we have immediately $\neg\mathrm{TT}t \rightarrow (\mathrm{T}t \leftrightarrow \mathrm{TT}t)$; for the other disjunct, $\neg\mathrm{TF}t \rightarrow \neg\mathrm{F}t \rightarrow \mathrm{T}t \rightarrow (\mathrm{TT} \rightarrow \mathrm{T}t)$. □

In fact, the well-known duality between consistency and completeness is preserved in the present setting, in that $\mathsf{CKF_{cs}}$ and $\mathsf{CKF_{cp}}$ are mutually $\mathcal{L}_{\mathbb{N}}$-interpretable via Cantini's dual translation, mapping T to $\neg$F [3]. More precisely, let c be a map of $\mathcal{L}_{\mathrm{T}}$ into itself preserving the arithmetical vocabulary, commuting with logical operations in the usual way, and mapping $\mathrm{T}x$ to $\neg\mathrm{F}x$.

Proposition 35. $\mathsf{CKF_{cs}}$ *and* $\mathsf{CKF_{cp}}$ *are mutually* $\mathcal{L}_{\mathbb{N}}$*-interpretable via* c.

Proof Sketch. For R1 and T1–T15, it suffices to observe that, within both $\mathsf{CKF_{cs}}$ and $\mathsf{CKF_{cp}}$, their instances A are self-dual, in the sense that $A \leftrightarrow A^c$. For example, for T15 we have (tacitly using the fact that $t \in \mathrm{Cterm}$)

$$\begin{array}{lll} (\mathrm{TFT}t \leftrightarrow \mathrm{TF}t)^c & \text{iff } \neg\mathrm{F}\neg\mathrm{F}\neg\neg\mathrm{T}\neg t \leftrightarrow \neg\mathrm{T}\neg\neg\mathrm{F}\neg t & \text{def of } c \\ & \text{iff } \neg\mathrm{F}\neg\mathrm{FT}\neg t \leftrightarrow \neg\mathrm{TF}\neg t & \text{T9} \\ & \text{iff } \neg\mathrm{TFT}\neg t \leftrightarrow \neg\mathrm{TF}\neg t & \text{def of F, T9.} \end{array}$$

The last line is a counterpositive instance of T15.

Similarly for TDel and TRep, just note that $\mathrm{TDel}^c = \neg\mathrm{F}\neg\mathrm{F}t \rightarrow \neg\mathrm{F}t \leftrightarrow (\neg\mathrm{TF}t \rightarrow \neg\mathrm{F}t)$, which is the counterpositive of TRep, and $\mathrm{TRep}^c = \neg\mathrm{F}t \rightarrow \neg\mathrm{F}\neg\mathrm{F}t \leftrightarrow (\neg\mathrm{F}t \rightarrow \neg\mathrm{TF}t)$, which is the counterpositive of TDel. □

The duality between $\mathsf{KF} + \mathsf{CONS}$ and $\mathsf{KF} + \mathsf{COMP}$ can be lifted to intertranslatability a.k.a. synonymy (see [19]); the same holds for $\mathsf{CKF_{cs}}$ and $\mathsf{CKF_{cp}}$.

Corollary 36. $\mathsf{CKF_{cs}}$ *and* $\mathsf{CKF_{cp}}$ *are synonymous.*

In particular, this means that the logical strength of $\mathsf{CKF_{cp}}$, too, coincides with that of CD^+.

6.2. Symmetric models. Combining the results on consistent and complete fixed points, it can also be shown that the definition of D can be adapted to obtain that the class of *symmetric* (i.e., consistent or complete) fixed points of $\mathscr{K}$ satisfies the axioms of $\mathsf{CD}^+_{\mathrm{T}}$ with D defined disjunctively as follows.

Definition 37 ($\mathsf{CD}^+_{\mathrm{T}}[\mathsf{SYM}]$). *Let* $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{SYM}]$ *be the* $\mathcal{L}_{\mathrm{T}}$*-theory whose axioms are those of* CD^+, *but where D defined in terms of* T *as*

$$\mathrm{D}x :\leftrightarrow (\mathrm{TT}x \vee \mathrm{TF}x) \wedge (\neg\mathrm{TT}x \vee \neg\mathrm{TF}x). \tag{15}$$

Using Facts 9 and 31, it can then be shown that symmetric fixed-point X can be used to obtain models of $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{SYM}]$.

Proposition 38. *For any symmetric fixed-point* X, $(\mathbb{N}, \mathbb{T}_X) \vDash \mathsf{CD}^+_{\mathrm{T}}[\mathsf{SYM}]$.

Proof. By induction on the length of proofs in $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{SYM}]$. For T3, one assumes $(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TT}\varphi \vee \mathrm{TF}\varphi \wedge (\neg\mathrm{TT}\varphi \vee \neg\mathrm{TF}\varphi)$ and may then reason by cases, depending on whether X is consistent or complete, following the arguments for Propositions 10 and 32. □

Accordingly, a corresponding CKF system is obtained by a disjunction of TDel and TRep.

Definition 39 ($\mathsf{CKF}_{\mathsf{sym}}$). $\mathsf{CKF}_{\mathsf{sym}}$ *is the system obtained from* $\mathsf{CKF}_{\mathsf{cs}}$ *by replacing TDel with*

$$(\mathrm{TT}t \to \mathrm{T}t^\circ) \vee (\mathrm{T}s^\circ \to \mathrm{TT}s). \tag{TSym}$$

To see that axiom TSym yields *consistency-or-completeness* $(\mathrm{TT}t \wedge \mathrm{TF}t) \to (\mathrm{TT}s \vee \mathrm{TF}s)$, assume $\mathrm{TT}t \wedge \mathrm{TF}t$. If $(\mathrm{TT}t \to \mathrm{T}t^\circ)$, then $\mathrm{T}t \wedge \mathrm{F}t$, which is impossible, hence $(\mathrm{T}s^\circ \to \mathrm{TT}s)$ and therefore $\neg\mathrm{TT}s \vee \neg\mathrm{TF}s$.

It can also be observed that the translations τ and ι defined in the proof of Proposition 22 yield mutual interpretability of $\mathsf{CKF}_{\mathsf{sym}}$ with $\mathsf{CT}[\mathsf{KF} + (\mathsf{CONS} \vee \mathsf{COMP})]$, hence we obtain the following equivalence.

Corollary 40. *The theories* CD^+, $\mathsf{CKF}_{\mathsf{cs}}$, $\mathsf{CKF}_{\mathsf{cp}}$, *and* $\mathsf{CKF}_{\mathsf{sym}}$ *have the same arithmetical consequences as the system* $\mathsf{RT}_{<\varepsilon_{\varepsilon_0}}$ *of ramified truth up to* $\varepsilon_{\varepsilon_0}$.

6.3. Mixed models. Mixed fixed-point models are those in which the truth predicate can feature *both* gaps and gluts. Can *mixed* (i.e., neither consistent nor complete) fixed-point model $\mathsf{CD}^+_{\mathrm{T}}$ under a suitable definition of D? We leave this question open. However, we observe that mixed fixed points are not models of $\mathsf{CD}^+_{\mathrm{T}}$ under any of the definitions of D considered above.

Proposition 41. *Let* X *be a mixed fixed point and let* $\mathsf{CD}^+_{\mathrm{T}}\star$ *range over the theories* $\mathsf{CD}^+_{\mathrm{T}}$, $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{COMP}]$, $\mathsf{CD}^+_{\mathrm{T}}[\mathsf{SYM}]$. *Then* $(\mathbb{N}, \mathbb{T}_X) \nvDash \mathsf{CD}^+_{\mathrm{T}}\star$.

Proof. In light of Proposition 10, we show that $(\mathbb{N}, \mathbb{T}_X) \nvDash$ T3 $\vee$ D5,[13] which are the two axioms where Facts 9 and 31 played a crucial role.

To show $(\mathbb{N}, \mathbb{T}_X) \nvDash$ T3, let λ and τ be such that $\lambda \wedge \neg\lambda \in X$ and $\tau \vee \neg\tau \notin X$, and assume moreover that $\mathsf{PAT} \vdash \lambda \leftrightarrow \neg\mathrm{T}\lambda$ and $\mathsf{PAT} \vdash \tau \leftrightarrow \mathrm{T}\tau$. Since $\neg(\tau \vee \lambda) \notin X$, it can be observed that the following jointly hold:

[13] This is of course redundant, but it clarifies the reasons why each of the axiom is not satisfied.

$$
\begin{aligned}
&(\mathbb{N}, X) \vDash \neg \mathrm{F}(\tau \vee \lambda), \\
&(\mathbb{N}, X) \vDash \mathrm{T}(\tau \vee \lambda), \\
&(\mathbb{N}, X) \vDash \neg(\tau \vee \lambda).
\end{aligned}
$$

We derive, for D defined by (15),

$$(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{D}(\tau \vee \lambda) \wedge \mathrm{TT}(\tau \vee \lambda) \wedge \neg \mathrm{T}(\tau \vee \lambda),$$

hence $(\mathbb{N}, \mathbb{T}_X) \nvDash$ T3.

To show $(\mathbb{N}, \mathbb{T}_X) \nvDash$ D5, let similarly φ be such that $\varphi \wedge \neg\varphi \in X$, and let ψ be such that $\psi \vee \neg\psi \notin X$. This entails

$$
\begin{aligned}
&(\mathbb{N}, X) \vDash \mathrm{F}(\varphi \wedge \psi) \wedge \neg \mathrm{T}(\varphi \wedge \psi), && \text{iff} \\
&(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{TF}(\varphi \wedge \psi) \wedge \neg \mathrm{TT}(\varphi \wedge \psi), && \text{hence} \\
&(\mathbb{N}, \mathbb{T}_X) \vDash \mathrm{D}(\varphi \wedge \psi).
\end{aligned}
$$

with D defined as per (15). However, $(\mathbb{N}, \mathbb{T}_X) \nvDash \mathrm{D}\varphi \vee \mathrm{D}\tau$ for any of the above definitions of D, since $(\mathbb{N}, \mathbb{T}_X) \vDash (\mathrm{TT}\varphi \wedge \mathrm{TF}\varphi) \wedge (\neg\mathrm{TT}\psi \wedge \neg\mathrm{TF}\psi)$, hence $(\mathbb{N}, \mathbb{T}_X) \nvDash$ D5. □

The reason why the class of symmetric, but not that of mixed, fixed points is suitable for modeling CD^+_T can be explained as follows. The former, but not the latter, features a specific interplay between the notions of *determinate* and *having a classical semantic value*. Within both the class of consistent and the class of complete fixed points, we can single out an intended interpretation, where being determinate can be defined as having a classical semantic value. These intended models are the least fixed point, and the largest fixed point, where the set of sentences with a classical value are those which are grounded in Kripke's sense. Proposition 41 clarifies why this is not possible in mixed models: in each of them, Boolean combinations of gluts and a gaps result in sentences which are *strictly* true or *strictly* false.[14]

§7. Assessment. The theories introduced in this work have unique features that place them among the most promising theories of truth available in the literature. In this section, we elaborate on some of these features, comparing variants of $\mathsf{CKF}_{\mathsf{cs}}$ with related approaches to truth and determinateness along the key dimensions outlined in the introductory section.

The generalizing function of truth, we argued, requires a strongly classical and fully compositional theory of truth. In our theories, such a function is realized in virtue of axioms T4–T6. For example, T4 is sufficient to exclude the existence of sentences that are both true and false and of sentences that are neither true nor false. In this sense, all variants of both CD and $\mathsf{CKF}_{\mathsf{cs}}$ are *classical theories of classical truth*. To state this formally, recall that the *internal theory* of a theory of truth S be defined as

$$\mathscr{I}(S) = \{A \mid S \vdash \mathrm{T}\ulcorner A \urcorner\}.$$

[14] This is essentially the same reason why the version of KF in [5] cannot satisfy the D-axioms of CD and CD^+ when

$$\mathrm{D}x :\leftrightarrow (\mathrm{T}x \vee \mathrm{F}x) \wedge \neg(\mathrm{T}x \wedge \mathrm{F}x).$$

For every variant CKF and CD, the *logic of* their internal theories is *classical*: (i) all (universal closures of) classical logical axioms in $\mathcal{L}_{\mathrm{T}}$ are true, and (ii) all classical logical inferences preserve truth, hence, (iii) all (universal closures of) theorems of classical logic in $\mathcal{L}_{\mathrm{T}}$ are true.[15]

By contrast, systems such as KF or DT are *classical theories of nonclassical truth*: they are formulated in classical logic, yet the logics governing their internal theories is nonclassical. As noted in the introduction, this compromises the generalizing power of its truth predicate. Advocates of such theories—e.g., [24]—stress that their internal theories, despite not obeying the laws of classical logic, enjoy other truth-theoretic virtues. For instance, the internal theory of $\mathsf{KF}+\mathsf{CONS}$ is not only closed under Strong-Kleene Logic, but it is *fully disquotational* too: $A \in \mathscr{I}(\mathsf{KF}+\mathsf{CONS})$ iff $\mathrm{T}\ulcorner A \urcorner \in \mathscr{I}(\mathsf{KF}+\mathsf{CONS})$. Full disquotation is often taken to be crucial by truth theorists (see, e.g., [7]). This property provably fails for the inner theory of any classical theory of classical truth that admits a standard model.

However, our results show that, in addition to the internal theory, another notion plays a prominent theoretical role in this context. Define the *deep theory* of S as

$$\mathscr{D}(S) = \{A \mid S \vdash \mathrm{T}\ulcorner \mathrm{T}\ulcorner A \urcorner\urcorner\}.$$

In theories of classical determinate truth such as the ones studied in this paper, the internal and deep theories *provably differ*.[16] One of the main virtues of the theories we propose is that the logics of their deep theories can be associated with well-known logics admitting a transparent truth predicate. For $\mathsf{CKF}_{\mathsf{cs}}$, the logic of its deep theory amounts to the familiar Strong-Kleene logic with a transparent truth predicate. Thus, although not every classical axiom will be in $\mathscr{I}(\mathsf{CKF}_{\mathsf{cs}})$, whatever *is* inside it can be closed under the relevant nonclassical rules of inferences and under iterations of T. Analogously, the logic of $\mathscr{D}(\mathsf{CKF}_{\mathsf{cp}})$ corresponds to the Logic of Paradox [21], and the logic of $\mathscr{I}(\mathsf{CKF}_{\mathsf{sym}})$ to Symmetric Strong-Kleene [2, 26].

As anticipated, our theories also provide clear semantic rules for the analysis of the language with type-free truth. To see this, we note that we can uniformly define a 'semantic' truth predicate $\mathrm{T}_{\mathrm{sem}}$ as $\mathrm{T}_{\mathrm{sem}}x :\leftrightarrow \mathrm{TT}x$. For such a predicate, the theories can prove universally quantified laws corresponding to the Strong-Kleene truth conditions for $\mathcal{L}_{\mathrm{T}}$.[17] In addition, the theories prove unrestricted positive and negative truth ascriptions for $\mathrm{T}_{\mathrm{sem}}$. In formalizing semantic rules for $\mathcal{L}_{\mathrm{T}}$, our theories also provide definite information on the space of 'models', or extensions of $\mathrm{T}_{\mathrm{sem}}$, that are admissible. While $\mathsf{CKF}_{\mathsf{cs}}$ only allows consistent interpretations of $\mathrm{T}_{\mathrm{sem}}$, $\mathsf{CKF}_{\mathsf{cp}}$ forces inconsistent but complete interpretations.

This is in stark contrast with CD and its variants, as they do not prove the Strong-Kleene conditions for $\mathrm{T}_{\mathrm{sem}}$ (cf. Corollaries 13 and 14). This means that, in such theories, truth behaves transparently and according to logical principles only on a restricted fragment of the language, namely, on determinate sentences. In addition, CD and

[15] For this induction to hold, it is important that the induction schema of $\mathsf{CKF}_{\mathsf{cs}}$ and variants is extended to T.

[16] For instance, in $\mathsf{CKF}_{\mathsf{cs}}$, $\lambda \vee \neg\lambda$ can be used to separate the two. Dually, one can use $\lambda \wedge \neg\lambda$ to separate the deep and the internal theory of $\mathsf{CKF}_{\mathsf{cp}}$, since the latter theory derives $\mathrm{TT}(\lambda \wedge \neg\lambda)$.

[17] Essentially, the compositional axioms of KF.

its variants do not impose clear conditions of admissible interpretations of the true and determinate sentences—the analogue of our $\mathrm{T}_{\mathrm{sem}}$. As the results in Fujimoto and Halbach show [13, theorems 7.11 and 8.1], the theories are compatible with consistent or inconsistent interpretation of the true and determinate sentences. In our approach, there is a clear choice to be made depending on one's chosen definition of determinateness: while the axioms of CD^+ formulated by means of the definition of D as $\mathrm{TT}x \vee \mathrm{TF}x$ result in the theory $\mathsf{CKF}_{\mathsf{cs}}$, the dual definition as $\neg\mathrm{TT}x \wedge \neg\mathrm{TF}x$ yields the theory $\mathsf{CKF}_{\mathsf{cp}}$ whose deep theory is inconsistent.

Combining these observations together, we see that variants of $\mathsf{CKF}_{\mathsf{cs}}$ (i) are fully compatible with the generalizing function of truth—unlike classical theories of nonclassical truth and (ii) capture a transparent and well-behaved notion of truth inside their deep theory—unlike CD and its variants.

Another fundamental feature of the theories introduced in this work is that the notion of determinateness, just like what happens in well-known formal approaches to truth, is *defined* in terms of truth. Feferman [6], for example, despite assigning priority to the axioms for determinateness over those for truth, defines a sentence to be determinate iff it is true or false (and not both). A similar case for a determinateness predicate defined in this way can be made for KF—see especially [5, 24].

Feferman's definition of determinateness is well suited for theories like DT and KF, which employ a self-applicable but nonclassical truth predicate. As Fujimoto and Halbach rightly point out, however, the definition $\mathrm{D}x :\leftrightarrow \mathrm{T}x \vee \mathrm{F}x$ is not appropriate for theories of a thoroughly classical conception of truth, such as CD and CD^+, or classical closures of Kripke–Feferman truth developed in this paper. This is why CD and CD^+ treat determinateness as a primitive.

By contrast, in our theories a strongly classical and compositional truth predicate coexists with a defined determinateness predicate. In particular, our results show that the desiderata imposed to the notion of truth and determinateness by theories such as CD^+ can in fact be realized by theories based on a defined determinateness predicate. While the definition $\mathrm{D}x :\leftrightarrow \mathrm{T}x \vee \mathrm{F}x$ is unsuitable in this context, the alternative $\mathrm{D}x :\leftrightarrow \mathrm{TT}x \vee \mathrm{TF}x$ is just right to license the principles of CD^+, and more generally to meet the core desiderata for a thoroughly classical, self-applicable conception of truth.

One might object that our definition of D seems more artificial than Feferman's, which rests on the natural thought that being determinate just means 'having a determinate (classical) truth value' in a paracomplete or paraconsistent model. But the definition we propose is in fact a rather natural incarnation of this standard notion. In particular, it puts $\mathsf{CKF}_{\mathsf{cs}}$'s determinateness in continuity with the notion of determinateness available within $\mathsf{KF} + \mathsf{CONS}$. Since in $\mathsf{KF} + \mathsf{CONS}$ determinateness is defined as $\mathrm{T}x \vee \mathrm{F}x$, extending the theory to its classical closure naturally requires introducing an additional layer of truth into the definition of D. If one endorses Feferman's extension of determinateness as given in various manifestations of the Kripke–Feferman theory, its extensions will remain unchanged in our theories. What changes is the generalizing power afforded by the classical truth-theoretic layer. More precisely, the $\mathbb{N}$-categoricity of our theories (Proposition 21) tells us that any ω-model of our theory features a standard, Kripke–Feferman notion of determinateness. And in each such model compositional axioms and classical logic are fully satisfied in a strong sense because it amounts to a classical closure of a Kripkean fixed point. In particular, if we consider the classical closure of the minimal fixed point of Kripke's theory of truth, the extension of its determinateness predicate is just the set of grounded sentences of $\mathcal{L}_{\mathrm{T}}$.[18]

We can reformulate the above observations without reference to Kripkean semantics. The Kripke–Feferman notion of determinateness amounts to being true or false. Its extension remains unchanged in (the consistent variant) of our theories, but its definition takes into account the strong classicality of our truth predicate required by its generalizing role. Determinateness is now defined as being *semantically true or false*, where 'semantically true' can be explicitly defined as the predicate $\mathrm{T}_{\mathrm{sem}}$ just introduced.

Taken together, the above observations point to a novel approach to truth and determinateness, yielding a family of axiomatic theories with distinctive features. Known theories that define determinateness in terms of truth and falsity—such as KF or DT—employ a compositional and self-applicable truth predicate, but one that, due to its nonclassical nature, does not perform well on generalizations such as the ones required in blind deductions.[19] Conversely, theories of classical truth like the Tarskian theory CT, variants of CD, or Friedman and Sheard's FS, feature a strongly compositional truth predicate, but one that is not suitable to capture the Kripke–Feferman notion of determinateness. In addition, for CT and FS, no equally satisfactory notion of determinateness is likely to be found.

However, the classical closures of KF show that it is possible to combine, in a single framework, both kinds of virtues: a classical, strongly compositional notion of truth that supports blind inferences, and the possibility of defining a class of determinate sentences satisfying desirable principles.[20]

§8. Extensions and open questions. To conclude, we list some open questions and potential lines of research stemming from our work.

(1) Each of the introduced variants of $\mathsf{CKF}_{\mathsf{cs}}$ restrict, in different ways, the class of Kripkean fixed points. For example, the axiom TDel in $\mathsf{CKF}_{\mathsf{cs}}$ restricts the fixed points to the consistent ones. Let $\mathsf{CKF} := \mathsf{CKF}_{\mathsf{cs}} - \mathrm{TDel}$. Then:
   (i) Is there a definition of D that ensures the identity of CKF and the appropriate reformulation of $\mathsf{CD}^{+}_{\mathrm{T}}$ in terms of such a definition?
   (ii) Would this definition of D ensure the soundness of $\mathsf{CD}^{+}_{\mathrm{T}}$ with respect to mixed models?

Moreover, one can ask whether there are natural principles that can be used to expand $\mathsf{CKF}_{\mathsf{cs}}$, and how strong the resulting theories become. We mention two possible such expansions.

(2) Given the semantics of $\mathsf{CKF}_{\mathsf{cs}}$, one might treat the least fixed point of $\mathscr{K}$ as the intended interpretation for its deep theory. One can then axiomatize this

[18] This makes fully transparent a link adumbrated by Fujimoto and Halbach, when they state: "call the sentences in *D determinate*. If it were not for the additional predicate D, they would be the sentences that are grounded in Kripke's [17] sense (with some qualifications)" [13, p. 222]. In our setting, this connection is made completely explicit without the need to reinterpret sentences featuring the primitive determinateness predicate.

[19] See again [12].

[20] As already mentioned in the introduction, a thorough philosophical assessment of the conceptual import of $\mathsf{CKF}_{\mathsf{cs}}$ will be carried out in future work. Our focus in this paper was on its formal properties.

conception by adding a minimality schema to $\mathsf{CKF_{cs}}$. So let CKF_μ be the theory obtained by expanding $\mathsf{CKF_{cs}}$ with the schema

$$K(\varphi(x)) \to \forall x(\mathrm{TT}x \to \varphi(x)),$$

where, for $\varphi \in \mathcal{L}_\mathrm{T}$, $K(\varphi(x))$ expresses that $\varphi(x)$ is a closed point of $\mathscr{K}$. Then:

(iii) What is the proof-theoretic strength of CKF_μ?

(3) One may also consider extensions of $\mathsf{CKF_{cs}}$ by means of reflection principles. It is well-known that the theory of truth enables one to directly express soundness extensions of theories in the form of *Global Reflection Principles* for a theory S: for any φ, if φ is provable in S, then $\mathrm{T}\varphi$. It is also well-known that the most prominent type-free theories of truth don't sit well with Global Reflection.[21] However, for such theories Global Reflection may not be the right soundness extension to focus on. Reinhardt [24] proposed to consider instead a partial reflection principle for *provably true* sentences. A similar strategy is available in $\mathsf{CKF_{cs}}$ and its variants, if one looks at theorems under two layers of truth. The principle

$$(\forall \varphi : \mathcal{L}_\mathrm{T})(\mathrm{Prov}_{\mathsf{CKF_{cs}}}(\mathrm{TT}\varphi) \to \mathrm{TT}\varphi) \tag{16}$$

is sound with respect to the semantics for $\mathsf{CKF_{cs}}$ and can be safely added (and iterated) over our theories. It can then be asked:

(iv) What is the strength of the reflection principle (16) and iterations thereof?[22]

(4) Finally, for technical interest, it remains to be explored whether an analogue of our results is available for the theory $\mathsf{CD_T}$—as well as its variant $\mathsf{CD_T[COMP]}$. Specifically:

(v) Does $\mathsf{CD_T}$ have a natural semantics validating the principle T2, i.e., $\mathrm{D}t \to \mathrm{TD}t$, but not necessarily its converse?

(vi) Can the semantic construction be axiomatized in such a way that the resulting theory is identical to $\mathsf{CD_T}$?

(vii) What is the proof-theoretic strength of the resulting systems?

Acknowledgments. We would like to thank Kentaro Fujimoto, Volker Halbach, Johannes Stern, and Bartosz Wcisło for discussing selected topics contained in the paper.

Funding. L.C. would like to acknowledge funding from the NCN, MAESTRO grant 2019/34/A/HS1/00399 'Epistemic and Semantic Commitments of Foundational Theories', the Italian Ministry of Education, University and Research, PRIN 2022 program 'HERB—Human Explanation of Robotic Behaviour' prot. 20224X95JC, and the Alexander von Humboldt Foundation, Humboldt Research Fellowship.

[21] For instance, the addition of the Global Reflection to KF results in internal inconsistency, while $\mathsf{KF} + \mathsf{CONS}$ and FS are outright inconsistent with Global Reflection. Relatedly, $\mathsf{CKF_{cs}}$ and CD are inconsistent with the ω-iterated Global Reflection Principle.

[22] A proof-theoretic analysis of reflection principles of this kind is currently being conducted in joint work of Luca Castaldo, Kentaro Fujimoto, and Maciej Głowacki.

C.N. acknowledges funding from the AHRC Arts and Humanities Research Council, grants AH/V015516/1 and AH/Z506515/1, and the EPSRC Engineering and Physical Sciences Research Council/EU Commission, grant 'PLEXUS: Philosophical, Logical, and Experimental Routes to Substructurality'.

BIBLIOGRAPHY

[1] Bacon, A. (2015). Can the classical logician avoid the revenge paradoxes? *Philosophical Review*, **124**(3), 299–352.

[2] Blamey, S. (2002). Partial logic. In Gabbay, D. M., and Guenther, F., editors. *Handbook of Philosophical Logic* (second edition), Vol. 5. Dordrecht: Kluwer, pp. 261–353.

[3] Cantini, A. (1989). Notes on formal theories of truth. *Zeitschrift für Logik un Grundlagen der Mathematik*, **35**, 97–130.

[4] ———. (1996). *Logical Frameworks for Truth and Abstraction: An Axiomatic Study*. Vol 135 of Studies in Logic and the Foundations of Mathematics. Amsterdam: Elsevier Science B.V.

[5] Feferman, S. (1991). Reflecting on incompleteness. *Journal of Symbolic Logic*, **56**, 1–49.

[6] ———. (2008). Axioms for determinateness and truth. *Review of Symbolic Logic*, **1**(2), 204–217.

[7] Field, H. (2008). *Saving Truth from Paradox*. Oxford: Oxford University Press.

[8] ———. (2022). The power of naive truth. *The Review of Symbolic Logic*, **15**(1), 225–258.

[9] Fischer, M., Halbach, V., Stern, J., & Jönne, K. (2015). Axiomatizing semantic theories of truth? *The Review of Symbolic Logic*, **8**(2), 257–278.

[10] Fischer, M., Nicolai, C., & Dopico, P. (2023). Nonclassical truth with classical strength. A proof-theoretic analysis of compositional truth over hype. *The Review of Symbolic Logic*, **16**(2), 425–448.

[11] Fujimoto, K. (2010). Relative truth definability of axiomatic truth theories. *Bulletin of Symbolic Logic*, **16**(3), 305–344.

[12] ———. (2022). The function of truth and the conservativeness argument. *Mind*, **131**(521), 129–157.

[13] Fujimoto, K., & Halbach, V. (2024). Classical determinate truth I. *The Journal of Symbolic Logic*, **89**(1), 218–261.

[14] Hájek, P., & Pudlák, P. (1998). *Metamathematics of First-Order Arithmetic*, Perspectives in Mathematical Logic. Berlin: Springer-Verlag. Second printing.

[15] Halbach, V., & Nicolai, C. (2018). On the costs of nonclassical logic. *Journal of Philosophical Logic*, **47**(2), 227–257.

[16] Halbach, V. (2014). *Axiomatic Theories of Truth* (Revised edition). Cambridge: Cambridge University Press, (first edition 2011).

[17] Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, **72**, 690–712.

[18] McGee, V. (1991). *Truth, Vagueness, and Paradox*. Indianapolis and Cambridge: Hackett Publishing.

[19] Nicolai, C. (2022). Gaps, gluts, and theoretical equivalence. *Synthese*, 200(366), 1–22.

[20] Pohlers, W. (1989). *Proof Theory: An introduction*. Berlin: Springer.

[21] Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, **8**(1), 219–241.

[22] Quine, W. V. O. (1990). *Pursuit of Truth*. Cambridge, Massachusetts: Harvard University Press.

[23] Reinhardt, W. (1985). Remarks on significance and meaningful applicability. *Mathematical Logic and Formal Systems*, **94**, 227.

[24] ———. (1986). Some remarks on extending and interpreting theories with a partial predicate for truth. *Journal of Philosophical Logic*, **15**, 219–251.

[25] Rogers, H. (1987). *Theory of Recursive Functions and Effective Computability*. New York: McGraw–Hill Book Company.

[26] Scott, D. (1975). Combinators and classes. In Böhm, C., editor. *International Symposium on Lambda-Calculus and Computer Science Theory*. Lecture Notes in Computer Science: Springer, pp. 1–26.

[27] Visser, A. (2006). Categories of theories and interpretations. In Enayat, A., Kalantari, I., and Moniri, M., editors. *Logic in Tehran*, Vol. 26, Lecture Notes in Logic. La Jolla: Association for Symbolic Logic, pp. 284–341.

MUNICH CENTER FOR MATHEMATICAL PHILOSOPHY
LMU MUNICH
MUNICH
GERMANY
E-mail: castaldluca@gmail.com

KING'S COLLEGE LONDON
UNITED KINGDOM
E-mail: carlo.nicolai@kcl.ac.uk