Metainferences from a Proof-Theoretic Perspective, and a Hierarchy of Validity Predicates

Rea Golan[1]

Abstract

I explore, from a proof-theoretic perspective, the hierarchy of classical and paraconsistent logics introduced by Barrio, Pailos and Szmuc in (*Journal o f Philosophical Logic*, *49*, 93-120, 2021). First, I provide sequent rules and axioms for all the logics in the hierarchy, for all inferential levels, and establish soundness and completeness results. Second, I show how to extend those systems with a corresponding hierarchy of validity predicates, each one of which is meant to capture "validity" at a different inferential level. Then, I point out two potential philosophical implications of these results. (i) Since the logics in the hierarchy differ from one another on the rules, I argue that each such logic maintains its own distinct identity (contrary to arguments like the one given by Dicher and Paoli in 2019). (ii) Each validity predicate need not capture "validity" at more than one metainferential level. Hence, there are reasons to deny the thesis (put forward in Barrio, E., Rosenblatt, L. & Tajer, D. (*Synthese*, 2016)) that the validity predicate introduced in by Beall and Murzi in (*Journal o f Philosophy*, *110*(3), 143–165, 2013) has to express facts not only about what follows from what, but also about the metarules, etc.

Keywords Metainferences · Validity predicate · Proof theory

✉   Rea Golan
     reagolan@gmail.com

1   Institut für Philosophie, Freie Universität, Habelschwerdter Allee 30, 14195, Berlin, Germany

1 Introduction

The question of what identifies a logic as such has recently attracted some attention in the discussion of substructural logics, in particular, in the discussion of the non-transitive strict-tolerant logic ST and its relation with classical logic (CL). Some proponents of ST (as notably stated by Ripley, e.g., in [16]) claim that it is identical with CL, because the two logics have the same consequence relation. Yet, unlike CL, ST can be conservatively extended with a transparent truth predicate, as well as with a validity predicate. Thus, other authors, such as Dicher and Paoli in [9], identify ST with the paraconsistent logic LP, and still others maintain it is different both from CL and from LP.

An impressive recent work by Barrio, Pailos, and Szmuc (BPS) [4] generalizes the ST phenomenon. BPS construct a hierarchy of infinitely many logics where each of those logics "agrees" with CL at more inferential levels than its predecessors, but deviates from CL from some level onwards. The first logic in the hierarchy is LP, which agrees with CL on theorems but doesn't approve some of the classical inferences like "explosion"; the second is ST, which agrees with CL on theorems and inferences, but differs on metainferences such as Cut; the third logic agrees on theorems, inferences, and metainferences, but differs on metametainferences, and so on. Based on this construction, BPS draw the philosophical conclusion that

> "the hierarchy allows to convey a sort of philosophical impossibility result: that it is impossible to identify a logic with its given set of valid inferences of any inferential level—for it is possible to find logics which will coincide in that respect, despite of being intuitively different from it... Indeed, the logics in our hierarchy show that it is equally dissatisfying to identify a logic with a set of valid metainferences...[O]r to identify a logic with its set of valid metametainferences, or with its set of valid metainferences of level 3, or 4, ..., or any arbitrarily large level. Whence, the need of an appropriate identity criterion for logical systems that can cope with cases of this sort". ([4, p. 20])

Thus, it seems that all inferential levels have to be taken into account in identifying a logic. In particular, ST cannot be identified with CL, for there are infinitely many inferential levels where they differ. In what follows, I wish to elaborate this line of thought. To that end, I shall present a proof-theoretic account of the BPS hierarchy, which they construct model-theoretically. This account may put us in a better position to address the question of whether some of these logics are identical with others, as well as related issues.

Here is how I proceed. In Section 2, I provide a brief introduction of LP, ST, and the debate on ST's identity. Then, I introduce the concept of *metainference*, which is central to this discussion, and review the BPS hierarchy. In Section 3, I provide proof-theoretic accounts for CL, LP, ST, and the rest of the logics in the BPS hierarchy, at all inferential levels, and establish soundness and completeness

results. In Section 4, I define a sequence of validity predicates, each one of which is meant to capture "validity" at a different inferential level. In Section 5, I point out two philosophical implications of these results. (i) Since the logics in the hierarchy differ from one another with respect to the rules they approve, each such logic maintains its own distinct identity. (ii) Each validity predicate need not capture "validity" at more than one metainferential level. Thus, I deny the thesis (put forward in [3]) that the validity predicate introduced in by Beall and Murzi in [6] has to express facts not only about what follows from what, but also about the metarules, etc.

N.B. Most of the proofs in the paper contain tedious details that can easily be worked out. For the sake of brevity I leave some of those details as an exercise. Also, to make the paper easier for reading, I put most of the proofs in the Appendix. Nonetheless, the proofs of the central results (and the non-trivial ones) are left in the body of the paper.

2 LP, ST, and the ST Hierarchy

Except for CL (with which I assume the reader is familiar) the logics to be considered here are all based on *strong Kleene valuations*, defined as follows. Let $\mathscr{L}$ be a formal language with the connectives $\neg, \vee, \wedge$. A valuation function $v : \mathscr{L} \to \{0, \frac{1}{2}, 1\}$ is *strong Kleene* if it is given by the truth tables:

$\neg$	
1	0
$\frac{1}{2}$	$\frac{1}{2}$
0	1

$\vee$	1	$\frac{1}{2}$	0
1	1	1	1
$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$
0	1	$\frac{1}{2}$	0

$\wedge$	1	$\frac{1}{2}$	0
1	1	$\frac{1}{2}$	0
$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0
0	0	0	0

The connectives $\supset, \leftrightarrow$ will be defined derivatively, in the usual way: $A \supset B \equiv \neg A \vee B$, $A \leftrightarrow B \equiv (A \supset B) \wedge (B \supset A)$.

LP is the logic that results from taking $\frac{1}{2}$ and 1 to be *designated values* [14]. That is, $\Gamma \models_{LP} \Delta$ if for all v, if $v(\gamma) \in \{\frac{1}{2}, 1\}$ for all $\gamma \in \Gamma$, then there is some $\delta \in \Delta$ such that $v(\delta) \in \{\frac{1}{2}, 1\}$. By contrast, ST has no designated values. Rather, ST sets different standards for premises and conclusions: premises have to be "strictly satisfied" (assigned the value 1) whereas conclusions just have to be "tolerantly satisfied" (assigned either $\frac{1}{2}$ or 1). That is, $\Gamma \models_{ST} \Delta$ if, for all v, if $v(\gamma) = 1$ for all $\gamma \in \Gamma$, then there is some $\delta \in \Delta$ such that $v(\delta) \in \{\frac{1}{2}, 1\}$.

As for proof systems, I shall work with the following sequent calculi for LP and ST, given in [5] and [17], respectively. The premises and the conclusions are read as sets, *not* as multi-sets; hence, Contraction and Exchange are built-in properties. LP is given by:

Axioms:

$$\Gamma, A \Rightarrow A, \Delta$$
$$\Gamma \Rightarrow A, \neg A, \Delta$$

Structural Rules:

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma, \Sigma \Rightarrow \Delta, \Pi} \; Weakening$$

$$\frac{\Gamma, A \Rightarrow \Delta \quad \Gamma \Rightarrow A, \Delta}{\Gamma \Rightarrow \Delta} \; Cut$$

Operational Rules:

$$\frac{\Gamma, A \Rightarrow \Delta \quad \Gamma, B \Rightarrow \Delta}{\Gamma, A \vee B \Rightarrow \Delta} \; L\vee \qquad \frac{\Gamma \Rightarrow A, B, \Delta}{\Gamma \Rightarrow A \vee B, \Delta} \; R\vee$$

$$\frac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \wedge B \Rightarrow \Delta} \; L\wedge \qquad \frac{\Gamma \Rightarrow A, \Delta \quad \Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \wedge B, \Delta} \; R\wedge$$

$$\frac{\Gamma, \neg A \vee \neg B \Rightarrow \Delta}{\Gamma, \neg(A \wedge B) \Rightarrow \Delta} \; \neg\wedge Left \qquad \frac{\Gamma \Rightarrow \neg A \vee \neg B, \Delta}{\Gamma \Rightarrow \neg(A \wedge B), \Delta} \; \neg\wedge Right$$

$$\frac{\Gamma, \neg A \wedge \neg B \Rightarrow \Delta}{\Gamma, \neg(A \vee B) \Rightarrow \Delta} \; \neg\vee Left \qquad \frac{\Gamma \Rightarrow \neg A \wedge \neg B, \Delta}{\Gamma \Rightarrow \neg(A \vee B), \Delta} \; \neg\vee Right$$

$$\frac{\Gamma, A \Rightarrow \Delta}{\Gamma, \neg\neg A \Rightarrow \Delta} \; \neg\neg L \qquad \frac{\Gamma \Rightarrow A, \Delta}{\Gamma \Rightarrow \neg\neg A, \Delta} \; \neg\neg R$$

and ST is given by:

Axioms:

$$\Gamma, A \Rightarrow A, \Delta$$

Structural Rules:

$$\frac{\Gamma \Rightarrow \Delta}{\Gamma, \Sigma \Rightarrow \Delta, \Pi} \; Weakening$$

Operational Rules:

$$\frac{\Gamma \Rightarrow A, \Delta}{\Gamma, \neg A \Rightarrow \Delta} \; L\neg \qquad \frac{\Gamma, A \Rightarrow \Delta}{\Gamma \Rightarrow \neg A, \Delta} \; R\neg$$

$$\frac{\Gamma, A \Rightarrow \Delta \quad \Gamma, B \Rightarrow \Delta}{\Gamma, A \vee B \Rightarrow \Delta} \; L\vee \qquad \frac{\Gamma \Rightarrow A, B, \Delta}{\Gamma \Rightarrow A \vee B, \Delta} \; R\vee$$

$$\frac{\Gamma, A, B \Rightarrow \Delta}{\Gamma, A \wedge B \Rightarrow \Delta} \; L\wedge \qquad \frac{\Gamma \Rightarrow A, \Delta \quad \Gamma \Rightarrow B, \Delta}{\Gamma \Rightarrow A \wedge B, \Delta} \; R\wedge$$

LP and ST are sound and complete with respect to the above calculi: for all $\Gamma, \Delta \subseteq \mathscr{L}$, $\Gamma \models_{LP} \Delta$ iff the sequent $\Gamma \Rightarrow \Delta$ is derivable in LP, and $\Gamma \models_{ST} \Delta$ iff the sequent $\Gamma \Rightarrow \Delta$ is derivable in ST. Note that ST is formulated with double-line operational rules, where rules with multiple premises ($L\vee$ and $R\wedge$) are "read" in

reverse as licensing the derivation of each premise from the conclusion.[1] On the other hand, LP has Cut, and so invertibility can be proven for its operational rules. Thus, there is no need to formulate these rules with double lines.

As is well known, ST is actually CL without Cut, but Cut is admissible as guaranteed by Gentzen's *Hauptsatz*. Therefore, CL and ST have the same the consequence relation; i.e., they are in full agreement on the status of inferences. Yet, CL and ST differ at a higher level, that of *metainferences*.

Since the concept of *metainference* is going to be at the center of my discussion, a few words about it are in order. Briefly put, metainferences are inferences between inferences (like metarules such as Cut), but also inferences between previously defined metainferences, e.g., an inference between previously defined metainferences $\dfrac{\Gamma_1 \Rightarrow \Delta_1}{\Gamma_2 \Rightarrow \Delta_2}$, $\dfrac{\Gamma_3 \Rightarrow \Delta_3}{\Gamma_4 \Rightarrow \Delta_4}$ and a previously defined metainference $\dfrac{\Gamma_5 \Rightarrow \Delta_5}{\Gamma_6 \Rightarrow \Delta_6}$:

$$\frac{\dfrac{\Gamma_1 \Rightarrow \Delta_1}{\Gamma_2 \Rightarrow \Delta_2} \qquad \dfrac{\Gamma_3 \Rightarrow \Delta_3}{\Gamma_4 \Rightarrow \Delta_4}}{\dfrac{\Gamma_5 \Rightarrow \Delta_5}{\Gamma_6 \Rightarrow \Delta_6}}$$

and so on. To be precise, metainferences are defined recursively as follows:

Definition 2.1 Let $SEQ^0(\mathscr{L})$ be the set of all inferences in the language $\mathscr{L}$, namely, sequents of the form $\Gamma \Rightarrow \Delta$ where $\Gamma, \Delta \subseteq \mathscr{L}$. For all $n \in \mathbb{N}$, let $SEQ^n(\mathscr{L})$ be the set of metainferences of level n, namely, sequents of the form $\Gamma \Rightarrow_n A$ where $\Gamma \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$. The members of Γ are called the *premise metainferences* and A is the *conclusion metainference*.

For the sake of clarity, I shall sometimes denote a metainference of level n, $\Gamma \Rightarrow_n A$ (where $\Gamma = \{\gamma_1, ..., \gamma_m\} \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$) also by $_{(n)} \dfrac{\gamma_1, ..., \gamma_m}{A}$, where the level of that metainference may be optionally mentioned inside brackets to the left of the horizontal proof line.[2] Following [9] and [4], I shall consider metainferences to be syntactic objects that form a genuine part of the system under consideration, or, as these authors put it, *bona fide* syntactic objects [9, p. 8];[4, p. 7].

From a model-theoretic point of view, valuation functions satisfy not only formulas and inferences, but also metainferences of all levels. The definition is straightforward, by way of generalization: we say that a valuation function v satisfies a metainference of level n $\Gamma \Rightarrow_n A$ ($v \models \Gamma \Rightarrow_n A$) if either v dissatisfies some $\gamma \in \Gamma$, or v satisfies A. This notion of satisfaction is sometimes called the "local"

---

[1]Importantly, the invertible rules make many metarules derivable rather than merely admissible. As Dicher and Paoli show in [9], this version of ST is sound and complete not only at the first inferential level, but also with respect to metainferences of level 1. See Section 3.3 below.

[2]I shall use this notation mainly in proof trees where metainferences of different levels are involved. In particular, I shall use the notation to indicate the degree and role of each horizontal line in those proof trees, i.e., whether it is a part of a "sequent," or whether it indicates a transition between "sequents". See Sections 3 and 4 below.

notion of metainferential validity, and I shall adhere to it. Thus, we may say that $\Gamma \Rightarrow_n A$ is "locally valid," or simply "valid" ($\models \Gamma \Rightarrow_n A$) if for all $v\colon v \models \Gamma \Rightarrow_n A$.[3]

At this point, there arises the question of whether ST and CL are actually identical, given that they differ somewhat abstractly, only at the level of metainferences. Needless to say, if the "identity" of a logic is determined solely in terms of its consequence relation, we must answer this question positively. Yet against such identification speaks the fact that unlike CL, ST can be conservatively extended with a transparent truth predicate and with a validity predicate [8, 16, 17].

Suppose then, that we refrain from identifying ST with CL. It would still be interesting to ask whether ST can be identified with some other well known logic. The following definition of *external logic* may be the key to answering this latter question in the affirmative.

Definition 2.2 Let L be some logic. The external logic of L is the logic that governs its theorems. Namely, Δ is derivable from Γ in the external logic of L iff the sequent $\Rightarrow \Delta$ is derivable from the sequents $\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m$ in L (where $\Gamma = \{\gamma_1, ..., \gamma_m\}$).

Interestingly, CL and LP each coincide with their own external logic, but ST doesn't [2]. For example, the consequence relation of ST is classical, so it has explosion: you can derive the sequent $p \wedge \neg p \Rightarrow q$ in ST. On the other hand, it is impossible to derive $\Rightarrow q$ from $\Rightarrow p \wedge \neg p$ in ST, the valuation $v(q) = 0$, $v(p) = \frac{1}{2}$ being a counterexample. In [2] Barrio, Rosenblatt, and Tajer observe that the external logic of ST is actually LP. To understand this better, consider the following translation function $lower : \bigcup_{n<\omega} SEQ^n(\mathscr{L}) \to \mathscr{L} \cup \bigcup_{n<\omega} SEQ^n(\mathscr{L})$, defined recursively by

$$lower(\Gamma \Rightarrow \Delta) = ((\wedge\Gamma) \supset (\vee\Delta))$$
$$lower(\Gamma \Rightarrow_{n+1} A) = \{lower(\gamma) \mid \gamma \in \Gamma\} \Rightarrow_n lower(A).$$

What Barrio, Rosenblatt, and Tajer prove is that $\Rightarrow \Delta$ is derivable from $\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m$ (where $\Gamma = \{\gamma_1, ..., \gamma_m\}$) in ST iff the sequent $\Gamma \Rightarrow \Delta$ is derivable in LP. In fact, they prove a stronger result (see Theorem 2.7 below), namely, that for all $n \in \mathbb{N} \cup \{0\}$: $\models_{ST} \Gamma \Rightarrow_n A$ iff $\models_{LP} lower(\Gamma \Rightarrow_n A)$. Based on these results, one could argue that ST is just LP in disguise, namely, LP in a different mode of presentation. Indeed, such an approach is advocated by Dicher and Paoli in [9].[4]

The last option I shall mention here is that ST is neither CL nor LP. Rather, ST is a member in an infinite hierarchy of logics that are all partly paraconsistent, yet progressively classical: the higher you climb up the hierarchy, the more "classical"

---

[3]It is worth mentioning that there is an alternative, "global" notion of validity for metainferences. According to this latter notion, $\Gamma \Rightarrow_n A$ is valid if, whenever the premise metainferences are satisfied by any valuation, so is the conclusion metainference. Yet, the global notion will not be considered here any further. See [4, 9] for discussions of why the local notion is preferable to the global one.

[4]Their argument is actually more complex, and relies both on the completeness of ST with respects to its metainferences (regarding the above formulation, with double-line rules) and on the claim that two logics are identical if they have the same "abstract consequence relation" as defined by Block and Jónsson, e.g., in [7]. Unfortunately, an elaborate discussion of this approach would exceed the scope of the present paper. See my discussion in Section 5.1 below.

the logics you come across. In brief, the idea is this: we saw that ST sets different standards for premises and conclusions: $\Gamma \Rightarrow \Delta$ is ST-valid if, given that its premises are K_3-valid—namely, they get the designated value of Kleene's K_3 (i.e., 1)—it cannot be the case that all of its conclusions are LP-invalid—namely, don't get one of LP's designated values: either $\frac{1}{2}$ or 1. Let us designate this fact by the notation ST= K_3/LP. In general, given two logics L_1, L_2, let us define the logic L_1/L_2 as the logic whose premise-standard is given by L_1 and conclusion-standard by L_2. Thus, BPS construct their hierarchy in the following way [4, p. 15]:

Definition 2.3 The collection $\mathbb{ST} = \{L_i | i \in \mathbb{N}\}$ of logical systems is recursively defined so that $L_0 = LP$, $L_1 = ST$, and for $j \geq 2$, $L_j = \overline{L_{j-1}}/L_{j-1}$ (where $\overline{L_j} = L_n/L_m$ if $L_j = L_m/L_n$).

Definition 2.4 For $j \geq 2$ and $L_j \in \mathbb{ST}$ and a metainference $\Gamma \Rightarrow_{j-1} A$, where $\Gamma \subseteq SEQ^{(j-2)}(\mathscr{L})$, $A \in SEQ^{(j-2)}(\mathscr{L})$, a valuation v satisfies $\Gamma \Rightarrow_{j-1} A$ in L_j ($v \models_{L_j} \Gamma \Rightarrow_{j-1} A$) if $v \nvDash_{\overline{L_{j-1}}} \gamma$ for some $\gamma \in \Gamma$, or $v \models_{L_{j-1}} A$. [4, p. 15]

Definition 2.4 tells us how metainferences of various levels are evaluated by each logic in the hierarchy. For each n, the logic L_n sets different standards for premises and conclusions of metainferences of level $n - 1$: $v \models_{L_n} \Gamma \Rightarrow_{n-1} A$ if $v \nvDash_{\overline{L_{n-1}}} \gamma$ for some $\gamma \in \Gamma$, or $v \models_{L_{n-1}} A$. For $j \geq n$, however, metainferences are evaluated by L_n uniformly, i.e., $v \models_{L_n} \Gamma \Rightarrow_j A$ if either $v \nvDash_{L_n} \gamma$ for some $\gamma \in \Gamma$, or $v \models_{L_n} A$. Likewise, for $j < n - 1$, a metainference $\Gamma \Rightarrow_j A$ is evaluated uniformly by L_n, namely, by the standards of L_{n-1}. Thus, the logics in the hierarchy basically differ on the level at which the standard for premise satisfatction and the standard for conclusion satisfaction diverge: the higher you climb up the hierarchy, the higher this level is.

To understand the hierarchy better, consider the following three theorems of BPS:

Lemma 2.5 *For all $j > n \geq 1$, for all $\Gamma \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$, for every v:*
$v \models_{L_{n+1}} \Gamma \Rightarrow_n A$ iff $v \models_{L_j} \Gamma \Rightarrow_n A$.

Theorem 2.6 *For all $n \geq 1$, for all $\Gamma \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$:*
$\models_{L_{n+1}} \Gamma \Rightarrow_n A$ iff $\models_{CL} \Gamma \Rightarrow_n A$. [4, p.17]

Theorem 2.7 *For all j, for all $n \geq 1$, for all $\Gamma \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$:*
$\models_{L_{j+1}} \Gamma \Rightarrow_n A$ iff $\models_{L_j} lower(\Gamma \Rightarrow_n A)$
In other words,
$\models_{L_{j+1}} \Gamma \Rightarrow_n A$ iff $\models_{LP} \underbrace{lower...(lower...}_{n\ times} (\Gamma \Rightarrow_n A)...)$. [4, p.19]

Lemma 2.5 states that, given a metainference of level n, all logics subscripted with index$\geq n + 1$ will agree on its status. On the other hand, Theorem 2.6 states that each L_n agrees with CL up to the $n - 1$ inferential level; i.e., L_n and CL agree on

metainferences of level $\leq n-1$. Theorem 2.7 states in general terms what we already observed about ST and LP: that *lower* serves as a translation function between each logic in the hierarchy and its predecessor.

An immediate result of these theorems, which will prove useful later on, is the following lemma:

Lemma 2.8 *For all $n \geq 1$, and for all $\Gamma \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$:*
$\models_{CL} \Gamma \Rightarrow_n A$ *iff* $\models_{CL} lower(\Gamma \Rightarrow_n A)$.

Note that CL and L_n differ at level n; i.e., they disagree on at least some metainferences of level n. I wish to conclude this section by pointing to a specific form of metainferences that will later have a role to play: $meta_n - Cut$, which is the metarule governing metainferences of level n (namely, a metainference of level $n+1$) defined as follows. First, I shall use the following notation, which will be of use throughout this paper. Let $\Gamma \Rightarrow_n A$ be an inference of level n. Let us denote by $\Gamma \Rightarrow_n A(\Sigma \Rightarrow \Pi)$ where $\Sigma, \Pi \subseteq \mathscr{L}$ the metainference that is identical to $\Gamma \Rightarrow_n A$ *except* for its root—the bottom sequent—which is replaced by $\Sigma \Rightarrow \Pi$. Likewise, we denote by $A(\Sigma \Rightarrow \Pi)$ the metainference that is identical to A *except* for its the bottom sequent, which is replaced by $\Sigma \Rightarrow \Pi$. For example, if Γ consists of the metainference $\dfrac{\Theta_1 \Rightarrow \Delta_1}{\Theta_2 \Rightarrow \Delta_2}$ and A is the metainference $\dfrac{\Phi_1 \Rightarrow \Psi_1}{\Phi_2 \Rightarrow \Psi_2}$ then $\Gamma \Rightarrow_2 A(\Sigma \Rightarrow \Pi)$ is the metainference:

$$(2) \quad \dfrac{(1)\ \dfrac{\Theta_1 \Rightarrow \Delta_1}{\Theta_2 \Rightarrow \Delta_2}}{(1)\ \dfrac{\Phi_1 \Rightarrow \Psi_1}{\Sigma \Rightarrow \Pi}}$$

and $A(\Sigma \Rightarrow \Pi)$ is simply $\dfrac{\Phi_1 \Rightarrow \Psi_1}{\Sigma \Rightarrow \Pi}$. (Trivially, if $\Sigma = \Phi_2$ and $\Pi = \Psi_2$, then $\Gamma \Rightarrow_2 A(\Sigma \Rightarrow \Pi)$ is just $\Gamma \Rightarrow_2 A$.) Using this notation, $meta_n - Cut$ can be defined as the following metainference (where $\Gamma \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$, $\Sigma, \Pi \subseteq \mathscr{L}$, and $B \in \mathscr{L}$):

$$(n+1) \quad \dfrac{\Gamma \Rightarrow_n A(\Sigma, B \Rightarrow \Pi) \quad \Gamma \Rightarrow_n A(\Sigma \Rightarrow B, \Pi)}{\Gamma \Rightarrow_n A(\Sigma \Rightarrow \Pi)} \ meta_n{-}Cut$$

Observe that $meta_0 - Cut$ is the well known Cut rule. The following proposition (proven by Pailos in [13]) shows that the logics in the BPS hierarchy progressively validate these rules.

Proposition 2.9 *For all $n \geq 1$, and for all $0 \leq k \leq n-1$: $\models_{CL} meta_k - Cut$, $\models_{L_{n+1}} meta_k - Cut$, but for all $k \geq n-1$: $\nvDash_{L_n} meta_k - Cut$.*

3 Sequent Rules and Axioms for Metainferences

To wrap things up, we saw that two logics in the hierarchy may agree at many inferential levels, and yet differ from one another from some level onwards. Therefore, to

identify a logic, all inferential levels need to be taken into account. This moral applies not only to the model-theoretic account of metainferences provided by BPS. Indeed, recall that metainferences are for us first and foremost *syntactic* objects, and so it is nothing but natural to provide a *proof − theoretic* account of them, one that would hopefully align with the model-theoretic account of local validity.[5] This section is dedicated to that task.

It should be clear by now that a proof-theoretic account of some logic in the hierarchy—undoubtedly, an account of some L_n for $n \geq 2$, but also accounts of CL, LP, and ST—cannot consist only of sequent rules and axioms governing the logic's inferences. Rather, we should provide sequent rules and axioms for metainferences of all levels. This will be done by way of generalization. In a "regular" sequent calculus, let us recall, the axioms take the form of schemata of inferences ("sequents") and the rules take the form of schemata of metainferences of level 1. Likewise, to determine the behavior of some logic at the n-th metainferential level, we need both axioms in the form of schemata of metainferences of level n, and rules in the form of schemata of metainferences of level $n + 1$. These axioms and rules will allow us to produce, for each n, proof trees with "sequents" in the form of metainferences of level n and "rules" in the form of metainferences of level $n + 1$. Moreover, each instance of a "rule" of level n will become an instance of an "axiom" that can serve as a "leaf" in proof trees of metainferences of level $n + 1$.

At this point, it is worthwhile to mention the distinction often made in proof theory between *operational* and *structural* rules. Roughly speaking, the former govern the inferential behavior of certain bits of the language—most notably, the logical constants—whereas the latter have a more general character. Obviously, a metarule that governs inferential behavior at a higher level may take one of these forms: it may be either a generalization of some operational rule, or a generalization of some structural rule. A question then arises as to whether a proof-theoretic account of metainferences requires both kinds of rules. One of the results I'm about to show is that at least for the propositional fragment, the metainferential behavior of each logic in the hierarchy is completely determined by higher level structural rules. In other words, there is no need to provide operational rules over and above the regular ones, i.e., those of level 1 that govern the sequents.

To start off, let us consider the translation function *higher*, which is the converse of *lower*.

Definition 3.1 $higher : \mathscr{L} \cup \bigcup_{n<\omega} SEQ^n(\mathscr{L}) \rightarrow \bigcup_{n<\omega} SEQ^n(\mathscr{L})$ is defined recursively, for $A \in \mathscr{L}$ by $higher(A) \Longrightarrow A$, and for each inferential level $n \in \mathbb{N} \cup \{0\}$ by

$$higher(\Gamma \Rightarrow \Delta) = \frac{\Rightarrow \gamma_1, \Rightarrow \gamma_2, ..., \Rightarrow \gamma_m}{\Rightarrow \Delta} \quad (where \ \Gamma = \{\gamma_1, \gamma_2..., \gamma_m\})$$

---

[5]It is worth mentioning that the account below doesn't aim to encompass the phenomenon known as "anti-validities". See [18] for a model-theoretic account of anti-validities in the BPS hierarchy. There is also much to be said about anti-validities in the BPS hierarchy from a proof-theoretic perspective, but such a project will have to wait for another occasion.

and for $n \geq 1$ by

$$higher(\Gamma \Rightarrow_n A) = \{higher(\gamma) \mid \gamma \in \Gamma\} \Rightarrow_{n+1} higher(A)$$

Note that the output of $higher$, when applied to a metainference of level n, is a metainference of level $n + 1$. As opposed to $lower$, though, $higher$ keeps all formulas unchanged. Note further that for all $n \in \mathbb{N} \cup \{0\}$: $lower(higher(\Gamma \Rightarrow_n A)) = \Gamma \Rightarrow_n A$. That is, $lower \circ higher = id_{\bigcup_{n<\omega} SEQ^n(\mathscr{L})}$. Unfortunately, the converse doesn't always hold: $lower(\Gamma \Rightarrow \Delta) = (\wedge\Gamma) \supset (\vee\Delta)$, but $higher((\wedge\Gamma) \supset (\vee\Delta) = \Rightarrow (\wedge\Gamma) \supset (\vee\Delta)$. Yet, an analogous requirement will be true of the following *structural* rules for metainferences of all levels (given in Definition 3.2), as stated in Proposition 3.3 (part (ii)) below:

Definition 3.2 The following is a list of structural rules for all metainferential levels[6] ($n \geq 1$), all of which are generalizations of the well-known structural rules.

1. Transitivity rules - $Tran_n = \{tran_n^p, tran_n^c\}$, governing the transitivity of premise metainferences and conclusion metainferences, respectively:

$$(n+1) \dfrac{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_{i-1} \Rightarrow_{n-1} A_{i-1}, \Gamma_i \Rightarrow_{n-1} A_i, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B} \quad (n) \dfrac{\Pi \Rightarrow_{n-1} C}{\Gamma_i \Rightarrow_{n-1} A_i}}{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_{i-1} \Rightarrow_{n-1} A_{i-1}, \Pi \Rightarrow_{n-1} C, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B}} \ Tran_n^p \quad (1 \leq i \leq k)$$

$$(n+1) \dfrac{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma_1 \Rightarrow_{n-1} B_1}, ..., (n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma_l \Rightarrow_{n-1} B_l} \quad (n) \dfrac{\Sigma_1 \Rightarrow_{n-1} B_1, ..., \Sigma_l \Rightarrow_{n-1} B_l}{\Pi \Rightarrow_{n-1} C}}{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Pi \Rightarrow_{n-1} C}} \ Tran_n^c$$

2. Monotonicity rules - $Weak_n = \{Weak_n^p \ Weak_n^b\}$, for weakening with premise metainferences and with the bottom sequent of a metainference, respectively:

$$(n+1) \dfrac{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B}}{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k, \Gamma_{k+1} \Rightarrow_{n-1} A_{k+1}}{\Sigma \Rightarrow_{n-1} B}} \ Weak_n^p$$

$$(n+1) \dfrac{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B(\Lambda \Rightarrow \Theta)}}{(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B(\Lambda, \Lambda' \Rightarrow \Theta, \Theta')}} \ Weak_n^b$$

3. Reflexivity axioms: $(n) \dfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_i \Rightarrow_{n-1} A_i, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Gamma_i \Rightarrow_{n-1} A_i} \ Ref_n$ ($1 \leq i \leq k$).

---

[6]Observe that the degree of each proof line is mentioned inside brackets to the left of the proof line. As I said in footnote 2, I shall use this notation mainly in proof trees, to indicate the degree and role of each horizontal line, i.e., whether it is a part of a "sequent," or whether it indicates a transition between "sequents".

4. A set Aux_n of auxiliary axioms of the form $^{(n)}\dfrac{\emptyset^{n-1}}{\Gamma \Rightarrow_{n-1} A}$ for all $\Gamma \subseteq SEQ^{n-2}(\mathscr{L})$, $A \in SEQ^{n-2}(\mathscr{L})$ such that $\Gamma \Rightarrow_{n-1} A$ is derivable via a proof tree of metainferences of level $n-1$ (and where $\emptyset^{n-1}$ is the empty metainference of level $n-1$, namely, the metainference of level $n-1$ whose premises and conclusions are all empty: $\emptyset^0 = \emptyset$, $\emptyset^1 = \frac{\emptyset}{\emptyset}$, $\emptyset^2 = \frac{\frac{\emptyset}{\emptyset}}{\frac{\emptyset}{\emptyset}}$, etc.).

Proposition 3.3 *Here are some useful facts about these axioms:*

(i)   *$Weak_n^b$ for $n \geq 1$ is redundant: given the well-known Weakening rule (for inferences), it follows by induction for all n, via $Tran_n^c$.*

(ii) *For each instance of such axioms/rules r of level n (except for $Weak_n^b$ which is redundant, given (i)), $lower(r)$, $higher(r)$ instantiate that same axiom/rule at levels $n-1, n+1$, respectively.*

(iii) *Given Cut, $meta_n$-Cut (for $n \geq 1$) is derivable via $Tran_n^c$.*

3.1 Soundness and Completeness for CL^∞

Definition 3.4 Let CL^0 be the proof system of classical logic (which is the same as that of ST given in Section 2, with the addition of Cut). We recursively define the sequent rules and axioms for metainferences of level $n+1$ to be

$$CL^{n+1} = CL^n \cup Tran_{n+1} \cup Weak_{n+1} \cup Ref_{n+1} \cup Aux_{n+1},$$

and $CL^\infty = \bigcup\limits_{n=1}^{\infty} CL^n$. We say that $\vdash_{CL^\infty} \Gamma \Rightarrow_n A$ iff there is some m such that $\vdash_{CL^m} \Gamma \Rightarrow_n A$; i.e., $\Gamma \Rightarrow_n A$ is derivable in CL^m. Observe that

$$CL^\infty = CL^0 \cup \bigcup\limits_{n=1}^{\infty} (Tran_n \cup Weak_n \cup Ref_n \cup Aux_n).$$

In other words, each CL^n consists of CL^0 along with all the structural rules up to level n.

Now, to prove completeness, we need the following proposition:

Proposition 3.5 *For all $n \in \mathbb{N} \cup \{0\}$, and for every metainference $\Gamma \Rightarrow_n A$: $\vdash_{CL^\infty} \Gamma \Rightarrow_n A$ iff $\vdash_{CL^\infty} lower(\Gamma \Rightarrow_n A)$.*

Theorem 3.6 *CL^∞ is sound and complete with respect to $\models_{CL}$: for all $n \in \mathbb{N} \cup \{0\}$ and for every metainference $\Gamma \Rightarrow_n A$, $\vdash_{CL^\infty} \Gamma \Rightarrow_n A$ iff $\models_{CL} \Gamma \Rightarrow_n A$.*

Proof By induction on n. The $n = 0$ case is just the soundness and completeness of CL^0. Assume that the claim holds for some n, and consider a metainference $\Gamma \Rightarrow_{n+1} A$ of level $n+1$. By Lemma 2.8, $\models_{CL} \Gamma \Rightarrow_{n+1} A$ iff $\models_{CL} lower(\Gamma \Rightarrow_{n+1} A)$. By the inductive hypothesis, this is the case iff $\vdash_{CL^\infty} lower(\Gamma \Rightarrow_{n+1} A)$, and by Proposition 3.5 this is the case iff $\vdash_{CL^\infty} \Gamma \Rightarrow_{n+1} A$. $\qquad\square$

3.2 Soundness and Completeness for LP^∞

Definition 3.7 Let LP^0 be the proof system of LP given in Section 2. We recursively define the sequent rules and axioms for metainferences of level n to be

$$LP^{n+1} = LP^n \cup Tran_{n+1} \cup Weak_{n+1} \cup Ref_{n+1} \cup Aux_{n+1},$$

and $LP^\infty = \bigcup_{n=1}^{\infty} LP^n$. We say that $\vdash_{LP^\infty} \Gamma \Rightarrow_n A$ iff there is some m such that $\vdash_{LP^m} \Gamma \Rightarrow_n A$; i.e., $\Gamma \Rightarrow_n A$ is derivable in LP^m. Observe that

$$LP^\infty = LP^0 \cup \bigcup_{n=1}^{\infty} (Tran_n \cup Weak_n \cup Ref_n \cup Aux_n).$$

Unfortunately, unlike in the CL case, it is *not* the case that $\models_{LP} \Gamma \Rightarrow_n A$ iff $\models_{LP} lower(\Gamma \Rightarrow_n A)$. In fact, even the deduction theorem fails to hold in LP.[7] As a result, there are no shortcuts in the completeness proof, and we need to work out the details with our own hands.

Theorem 3.8 *LP^∞ is sound and complete with respect to $\models_{LP}$: for all $n \in \mathbb{N} \cup \{0\}$, and for every metainference $\Gamma \Rightarrow_n A$, $\models_{LP} \Gamma \Rightarrow_n A$ iff $\vdash_{LP^\infty} \Gamma \Rightarrow_n A$.*

Proof By induction. The $n = 0$ case is the soundness and completeness of LP, as proven in [5], following [15, p. 157, Theorem 8.7.9]. Soundness is easy to prove for all levels, and so I shall focus on completeness. Specifically, I shall discuss in detail the $n = 1$ case, which is relatively non-abstract and easy to follow. Then I shall point out why the proof for all $n > 1$ is analogous, after due adjustments are made.

So, let $\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m, \Sigma \Rightarrow \Pi$ be some sequents such that for all v: $v \models_{LP} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi}$. We want to show that $\vdash_{LP^\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi}$. As usual in this kind of proofs, let us show instead that $\nvdash_{LP^\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi}$ entails $\nvDash_{LP^\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi}$. To this end, assume that $\nvdash_{LP^\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi}$. Let $\{A_n\}_{n<\omega}$ be an enumeration of all the formulas in the language, and let $\{\Sigma_n\}_{n<\omega}$ be defined as follows: $\Sigma_0 = \Sigma$, and

$$\Sigma_{n+1} = \begin{cases} \Sigma_n \cup \{A_n\} & if \ \nvdash_{LP^\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, A_n \Rightarrow \Pi} \\ \Sigma_n & otherwise. \end{cases}$$

We then define $\Sigma' = \bigcup_{n=0}^{\infty} \Sigma_n$.

---

[7]Moreover, LP has Cut, so $\models_{LP} \dfrac{A \Rightarrow B \quad B \Rightarrow C}{A \Rightarrow C}$, but $\nvDash_{LP} A \supset B, B \supset C \Rightarrow A \supset C$.

Now, our sequent rules are usually defined only where the contexts are finite sets, and Σ' may be infinite, but it is rather obvious that for all finite $\Lambda \subseteq \Sigma'$:

$$\nvdash_{LP\infty} \frac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Lambda \Rightarrow \Pi}.$$ Moreover, Σ' is clearly a maximal such set: for all

$A \notin \Sigma'$ there is some n such that $\vdash_{LP\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, A \Rightarrow \Pi}$. Additionally, I

claim that Σ' has the following properties:

(i) $A \wedge B \in \Sigma'$ iff $A, B \in \Sigma'$
(ii) $A \vee B \in \Sigma'$ iff either $A \in \Sigma'$ or $B \in \Sigma'$
(iii) For all $A \in \mathscr{L}$, either $A \in \Sigma'$ or $\neg A \in \Sigma'$ (This is *not* an exclusive disjunction).

(i) and (ii) follow immediately from the invertibility of the $\wedge$, $\vee$ rules,[8] along with one application of $Tran_1$. For instance, the right-to-left direction of (i) can be proven based on the fact that $L\wedge$ is invertible. That is, for all $n \in \mathbb{N}$:

$$(2) \quad \dfrac{(1) \dfrac{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, A \wedge B \Rightarrow \Pi} \quad (1) \dfrac{\Sigma_n, A \wedge B \Rightarrow \Pi}{\Sigma_n, A, B \Rightarrow \Pi} L\wedge \; (reversed)}{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m} Tran_1^c}{\Sigma_n, A, B \Rightarrow \Pi}$$

and so $A \wedge B \notin \Sigma'$ entails $A, B \notin \Sigma'$. As for (iii), suppose that there is some A such that $A, \neg A \notin \Sigma'$. By maximality, there is some k such that for all $n \geq k$:

$\vdash_{LP\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, A \Rightarrow \Pi}$ and $\vdash_{LP\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, \neg A \Rightarrow \Pi}$. But on the

other hand, $\vdash_{LP\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n \Rightarrow A, \neg A, \Pi}$ since the bottom sequent is an axiom.[9]

Two application of $meta_1 - Cut$ (which is available by Proposition 3.3) will then give us

$$\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, A \Rightarrow \Pi} \quad \dfrac{\dfrac{\dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, \neg A \Rightarrow \Pi} \quad \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n \Rightarrow A, \neg A, \Pi}}{\dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n \Rightarrow A, \Pi}} meta_1 - Cut}{\dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n \Rightarrow \Pi}} meta_1 - Cut$$

contradicting our assumption that the bottom metainference is unprovable. To sum up, Σ' has the properties (i)-(iii).

---

[8] As I said in Section 2, invertibility can be proven in the usual way, since LP has Cut.

[9] To be precise, the bottom sequent is an axiom, and so $\dfrac{\emptyset}{\Sigma_n \Rightarrow A, \neg A, \Pi}$ is an instance of Aux_1. There-fore, we can weaken m times (applying $Weak_1^P$) with $\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m$, to get the required result.

Now, let v be the valuation function defined over the atoms by:

$$v(p) = \begin{cases} 0 & p \notin \Sigma' \\ \frac{1}{2} & p, \neg p \in \Sigma' \\ 1 & p \in \Sigma', \neg p \notin \Sigma' \end{cases}$$

v can clearly be extended in a well defined way to all $A \in \mathcal{L}$. Moreover, as in the completeness proof for sequents, it can be shown that for all $A \in \mathcal{L}$: $v(A) \in \{\frac{1}{2}, 1\}$ iff $A \in \Sigma'$. As a result, for all $\pi \in \Pi$: $v(\pi) = 0$, because $\Sigma' \cap \Pi = \emptyset$. (Otherwise, there would be some n such that $\Sigma_n \Rightarrow \Pi$ holds).

It remains to show that v forms a counterexample model to our metainference. Let us assume, towards reaching a contradiction, that this is not the case. Clearly, v doesn't satisfy the bottom sequent: we saw that $v(\sigma) \in \{1, \frac{1}{2}\}$ for all $\sigma \in \Sigma$ and that $v(\pi) = 0$ for all $\pi \in \Pi$. So if v doesn't form a counterexample model to the metainference, it has to dissatisfy at least one of the premise sequents, say, $\Gamma_i \Rightarrow \Delta_i$ for some i. Without loss of generality, we can assume that Δ_i consists of only one formula (otherwise, just replace it with a large disjunction, as usual). Namely, our rebellious sequent is of the form $\Gamma_i \Rightarrow A$, and $v \nvDash_{LP} \Gamma_i \Rightarrow A$. Therefore, $v(\gamma) \in \{\frac{1}{2}, 1\}$ for all $\gamma \in \Gamma_i$, but $v(A) = 0$. By the definition of v, we get $A \notin \Sigma'$, and so, by maximality, we get that there is some $k \in \mathbb{N}$ such that for all $n \geq k$:
$\vdash_{LP\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma_i \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, A \Rightarrow \Pi}$. Therefore, we can reason as follows:

$$\begin{array}{c} (2) \dfrac{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma_i \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, A \Rightarrow \Pi}}{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma_i \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, \Gamma_i, A \Rightarrow \Pi}} Weak_1^b \quad (2) \dfrac{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma_i \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Gamma_i \Rightarrow A} Ref_1}{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma_i \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, \Gamma_i \Rightarrow A, \Pi}} Weak_1^b \\ \hline (2) \dfrac{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma_i \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, \Gamma_i \Rightarrow \Pi}}{} meta_1-Cut \end{array}$$

Hence, $\vdash_{LP\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, \Gamma_i \Rightarrow \Pi}$. However, $\Sigma_n \cup \Gamma_i \subseteq \Sigma'$, because for all $\gamma \in \Gamma_i$: $v(\gamma) \in \{\frac{1}{2}, 1\}$, and so, by the construction of Σ': $\nvDash_{LP\infty} \dfrac{\Gamma_1 \Rightarrow \Delta_1, ...\Gamma \Rightarrow A, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma_n, \Gamma_i \Rightarrow \Pi}$, and we get a contradiction.

Consequently, v has to either satisfy A or dissatisfy some $\gamma \in \Gamma_i$. That is, v has to satisfy all the premise sequents of $\dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi}$, although it dissatisfies its conclusion sequent. Thus, v forms the counterexample model we were looking for. This completes the completeness proof for metainferences of level 1.

As I said, the completeness proof for metainferences of level $n + 1$ on the basis of the completeness of metainferences of level n is analogous to the above proof. Let me briefly sketch the general idea. Given a metainference of level $n + 1$ that is not derivable: $\nvdash_{LP\infty} {}^{(n+1)} \dfrac{\Gamma_1 \Rightarrow_n \Delta_1, ..., \Gamma_m \Rightarrow_n \Delta_m}{\Sigma \Rightarrow_n \Pi}$, we note that $\nvdash_{LP\infty} \Sigma \Rightarrow_n \Pi$.[10] $\Sigma \Rightarrow_n \Pi$ is of the form $\Sigma \Rightarrow_n \Pi(\Lambda \Rightarrow \Theta)$ for some Λ, Θ, and so,

---

[10] See the previous footnote for an explanation.

as before, we construct a maximal $\Lambda' \supset \Lambda$ such that for all finite $\Lambda'' \subseteq \Lambda'$: $\nvdash_{LP\infty} {}^{(n+1)} \dfrac{\Gamma_1 \Rightarrow_n \Delta_1, ..., \Gamma_m \Rightarrow_n \Delta_m}{\Sigma \Rightarrow_n \Pi(\Lambda'' \Rightarrow \Theta)}$ and, in particular, $\nvdash_{LP\infty} \Sigma \Rightarrow_n \Pi(\Lambda'' \Rightarrow \Theta)$. Now, Λ' can be shown to have all the required properties (i)–(iii), so as in the $n = 1$ case, we define a valuation function v, based on Λ'. Since for all finite $\Lambda'' \subseteq \Lambda'$: $\Sigma \Rightarrow_n \Pi(\Lambda'' \Rightarrow \Theta)$ is an unprovable metainference of level n, it follows by the inductive hypothesis that v satisfies all of its components-sequents *except for* its bottom sequent. It remains to show that v satisfies all $\Gamma_1 \Rightarrow_n \Delta_1, ..., \Gamma_m \Rightarrow_n \Delta_m$. The latter claim can be proven in a way similar to the $n = 1$ case, which requires the applications of Ref_{n+1}, $Weak_{n+1}^b$, and $meta_{n+1} - Cut$. This completes the proof. $\qquad\square$

3.3 Soundness and Completeness for ST^∞

Definition 3.9 Let ST^0 be the proof system of ST, given in Section 2. We recursively define

$$ST^{n+1} = ST^n \cup Tran_{n+1} \cup Weak_{n+1} \cup Ref_{n+1} \cup Aux_{n+1}$$

and $ST^\infty = \bigcup\limits_{n=1}^{\infty} ST^n$. Observe that

$$ST^\infty = ST^0 \cup \bigcup\limits_{n=1}^{\infty} (Tran_n \cup Weak_n \cup Ref_n \cup Aux_n).$$

Theorem 3.10 $\vdash_{ST\infty}$ *is sound and complete with respect to* $\models_{ST}$: *for all* $n \in \mathbb{N} \cup \{0\}$ *and for every metainference* $\Gamma \Rightarrow_n A$, $\vdash_{ST\infty} \Gamma \Rightarrow_n A$ *iff* $\models_{ST} \Gamma \Rightarrow_n A$.

Proof The $n = 0$ case is just the soundness and completeness of ST, proven, e.g., in [17]. For $n = 1$, namely, soundness and completeness for metainferences of level 1, see [9]. For all $n \geq 2$, and for a metainference $\Gamma \Rightarrow_n A$, let us prove by induction that $\vdash_{ST\infty} \Gamma \Rightarrow_n A$ iff $\vdash_{LP\infty} lower(\Gamma \Rightarrow_n A)$. If so, then for all $n \geq 2$: $\vdash_{ST\infty} \Gamma \Rightarrow_n A$ iff $\vdash_{LP\infty} lower(\Gamma \Rightarrow_n A)$, and by Theorem 3.8 this is the case iff $\models_{LP} lower(\Gamma \Rightarrow_n A)$, which is equivalent, by Theorem 2.7 to $\models_{ST} \Gamma \Rightarrow_n A$. So we would get $\vdash_{ST\infty} \Gamma \Rightarrow_n A$ iff $\models_{ST} \Gamma \Rightarrow_n A$, as required.

The base case is $n = 2$. By Proposition 3.3, for each axiom/rule r of level 2, $lower(r)$ instantiates that same axiom/rule, and since ST^∞ and LP^∞ share the structural axioms/rules at levels ≥ 1, it follows that $\vdash_{ST\infty} \Gamma \Rightarrow_n A$ entails $\vdash_{LP\infty} lower(\Gamma \Rightarrow_n A)$. On the other hand, we need to prove that for each axiom/rule of LP^∞ of level 1, q, $higher(q)$ is derivable in ST^∞. That proof is carried out along the lines of the proof of Proposition 3.5. In fact, it is rather easy to show that for each such axiom/rule, q, $higher(q)$ is derivable in ST^∞, using $Tran_1^c$ and $Tran_1^p$. For

example, here is a proof that $higher(\neg \wedge Right)$ is derivable in ST^∞: observe that in ST, as in CL: $\dfrac{\Rightarrow \neg A \vee \neg B}{\Rightarrow \neg(A \wedge B)}$. Hence, we can reason as follows:

$$(2) \ \cfrac{(1) \ \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow \neg A \vee \neg B, \Delta} \quad (1) \ \cfrac{\Rightarrow \neg A \vee \neg B}{\Rightarrow \neg(A \wedge B)} \ Tran_1^c}{(1) \ \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow \neg(A \wedge B), \Delta}}$$

The cases of the other axioms/rules are analogous, so they are left to the reader.

It remains to make the inductive step. However, note that from level 2 onwards, ST^∞ and LP^∞ have the same axioms/rules, namely, only the structural ones. By Proposition 3.3, for each such axiom/rule r, $lower(r)$ and $higher(r)$ instantiate that same axiom/rule, at levels $n - 1, n + 1$, respectively. Hence, for $n \geq 3$, and for a metainference $\Gamma \Rightarrow_n A$, it is clear that $\vdash_{ST^\infty} \Gamma \Rightarrow_n A$ iff $\vdash_{LP^\infty} lower(\Gamma \Rightarrow_n A)$. $\square$

3.4 Soundness and Completeness for the Rest of the Logics

To provide sequent rules and axioms for the rest of the logics in the hierarchy, we first need to introduce additional structural rules for all inferential levels. These are the substitution rules $Sub_n = \{Sub_n^p, Sub_n^c\}$:

$$(n+1) \ \cfrac{(n) \ \cfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_{i-1} \Rightarrow_{n-1} A_{i-1}, \Gamma_i \Rightarrow_{n-1} A_i, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B} \quad (n) \ \cfrac{\Pi \Rightarrow_{n-1} C}{\Gamma_i \Rightarrow_{n-1} A_i}}{(n) \ \cfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_{i-1} \Rightarrow_{n-1} A_{i-1}, \Pi \Rightarrow_{n-1} C, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B}} \ Sub_n^p \quad (1 \leq i \leq k)$$

$$(n+1) \ \cfrac{(n) \ \cfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Sigma \Rightarrow_{n-1} B} \quad (n) \ \cfrac{\Sigma \Rightarrow_{n-1} B}{\Pi \Rightarrow_{n-1} C}}{(n) \ \cfrac{\Gamma_1 \Rightarrow_{n-1} A_1, ..., \Gamma_k \Rightarrow_{n-1} A_k}{\Pi \Rightarrow_{n-1} C}} \ Sub_n^c$$

These rules allow us to substitute premise/conclusion metainferences for *equivalent* metainferences: that is why each Sub rule has as a premise a double-line metainference, whose role is to guarantee that the substitute and the substitutee are indeed equivalent. These substitution rules are clearly derivable from $Tran_n^p, Tran_n^c$, but they are weaker than the transitivity rules. Indeed, it is exactly this double line requirement which makes the difference between the Sub rules and the $Tran$ rules. Thus, there is no way to derive $meta_n - Cut$ only with the substitution rules (recall Proposition 3.3). Yet, observe that like the other structural rules, for each substitution rule r of level n, $lower(r)$ and $higher(r)$ instantiate that same rule at levels $n - 1, n + 1$, respectively.

Now, these rules are required for proof-theoretic accounts of the rest of the logics in the hierarchy. To understand this point, recall that each such logic becomes non-transitive at the level where its standard for premise satisfaction and its standard for conclusion satisfaction diverge. For ST ($m = 1$) these standards diverge at level 0 (for premises and conclusions of inferences), and in general, each L_m ($m \geq 1$) becomes non-transitive at level $m - 1$. Consequently, each L_m doesn't admit $Tran_{m-1}$. On

the other hand, for $n \neq m - 1$, L_m sets the same standards for satisfying premises and conclusions, and so it does admit $Tran_n$.[11] Now, the Sub_n rules substitute for the absent $Tran_n$ in order to guarantee that metainferences remain intersubstitutable with equivalent metainferences, even if transitivity no longer holds.

It follows from the above that each logic L_m in the BPS hierarchy (when $m \geq 2$) is governed at the metainferential level n ($n \in \mathbb{N} \cup \{0\}$) by the following sequent rules and axioms:

$$L_m^n = \begin{cases} CL^n & 0 \leq n \leq m - 2 \\ L_m^{n-1} \cup Weak_n \cup Ref_n \cup Aux_n \cup Sub_n & n = m - 1 \\ L_m^{n-1} \cup Weak_n \cup Ref_n \cup Aux_n \cup Tran_n & n > m - 1. \end{cases}$$

Let us define $L_m^\infty = \bigcup_{n=1}^\infty L_m^n$. Observe that for each m, we have

$$L_m^\infty = CL^{m-2} \cup Weak_{m-1} \cup Ref_{m-1} \cup Aux_{m-1} \cup Sub_{m-1} \cup \bigcup_{n=m}^\infty (Weak_n \cup Ref_n \cup Aux_n \cup Sub_n).$$

Now we can prove the following results:

Theorem 3.11 *For all $m \geq 2$, for all $n \in \mathbb{N} \cup \{0\}$, and for every metainference $\Gamma \Rightarrow_n A$: $\vdash_{L_m^\infty} \Gamma \Rightarrow_n A$ iff $\vdash_{L_{m-1}^\infty} lower(\Gamma \Rightarrow_n A)$.*

Theorem 3.12 *For all $m \in \mathbb{N} \cup \{0\}$, for all $n \in \mathbb{N} \cup \{0\}$, and for every metainference $\Gamma \Rightarrow_n A$: $\vdash_{L_m^\infty} \Gamma \Rightarrow_n A$ iff $\models_{L_m} \Gamma \Rightarrow_n A$.*

Proof By induction. The base cases are those of LP^∞, ST^∞, which have already been proven. Assume that the claim holds for some $m \geq 1$, and let $\Gamma \Rightarrow_n A$ be a metainference of level n (for some n). Assume that $\vdash_{L_{m+1}^\infty} \Gamma \Rightarrow_n A$. By Theorem 3.11, this is the case iff $\vdash_{L_m^\infty} lower(\Gamma \Rightarrow_n A)$, and by the inductive hypothesis this is the case iff $\models_{L_m} lower(\Gamma \Rightarrow_n A)$. By Theorem 2.7, that is equivalent to $\models_{L_{m+1}} \Gamma \Rightarrow_n A$. To sum up, $\vdash_{L_m^\infty} \Gamma \Rightarrow_n A$ iff $\models_{L_m} \Gamma \Rightarrow_n A$. □

4 A Hierarchy of Validity Predicates

From the point of view of the hierarchy, "validity" should no longer be considered a unified phenomenon. Rather, we have to distinguish between different concepts of "validity," pertaining to different metainferential levels. Therefore, the thought goes, a definition of one such concept, say "validity" as pertaining to inferences (metainferences of level 0) need not, and may not capture "validity" as pertaining to other metainferential levels.

---

[11] However, the damage is already done, so to speak, at all other levels. Recall Proposition 3.3: we need $Tran_n$ to prove $meta_k - Cut$ for $k \geq n$. Thus, the absence of transitivity rules at some level makes it impossible to derive $meta_n - Cut$ from that level onwards.

Therefore, the hierarchy pushes us to consider the possibility of a sequence of "validity" concepts—ultimately, validity predicates—each one of which is supposed to capture "validity" at a different inferential level. This section deals with that thought, and the next section deals with some of its (potential) philosophical implications.

Assume that in our language we have names not only for propositions, but also for sets of propositions, as well as for sequents, metainferences, metametainferences, etc.[12] To put the goal in more concrete terms, what I wish to do is to define a sequence of validity predicates $< Val_n \mid n \in \mathbb{N} \cup \{0\} >$ such that, $Val_0(\langle \Gamma \rangle, \langle \Delta \rangle)$ holds iff the sequent $\Gamma \Rightarrow \Delta$ is valid,[13] $Val_1(\langle \Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k \rangle, \langle \Sigma \Rightarrow \Pi \rangle)$ holds iff

$$\frac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k}{\Sigma \Rightarrow \Pi}$$

is valid, $Val_2(\langle \frac{\Gamma_1 \Rightarrow \Delta_1}{\Gamma_2 \Rightarrow \Delta_2} \rangle, \langle \frac{\Gamma_3 \Rightarrow \Delta_3}{\Gamma_4 \Rightarrow \Delta_4} \rangle)$ holds iff

$$(2) \frac{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1}{\Gamma_2 \Rightarrow \Delta_2}}{(1) \dfrac{\Gamma_3 \Rightarrow \Delta_3}{\Gamma_4 \Rightarrow \Delta_4}}$$

is valid, and so on. In other words, it seems like a condition of adequacy on the definitions of these predicates that what Beall and Murzi call the $V - schema$ [6, p. 153] holds. For inferences, the schema goes as follows:

$$\Gamma \Rightarrow \Delta \; iff \; \Rightarrow Val(\langle \Gamma \rangle, \langle \Delta \rangle)$$

and in general, the requirement is that a Val_n statement holds $if \; and \; only \; if$ the corresponding metainference is valid.

How should we go about defining all these predicates? Fortunately, Beall and Murzi in [6] have already shown us how to inferentially define the first predicate;[14] I shall simply follow their footsteps in defining the other predicates by way of generalization. The first predicate, Val_0, is to be defined (according to Beall and Murzi) by the following rules:

$$\frac{}{\Gamma, Val_0(\langle \Gamma \rangle, \langle \Delta \rangle) \Rightarrow \Delta} \; V_0 D \qquad \frac{\Gamma \Rightarrow \Delta}{\Rightarrow Val_0(\langle \Gamma \rangle, \langle \Delta \rangle)} \; V_0 P$$

$V_0 P$ is an introduction rule. It allows us to introduce Val_0 on the right if the relevant sequent holds. The idea is that if an inference is valid, then "saying as much—using the validity predicate—is true in a validity-strength fashion" [6, p. 152]. In other

---

[12]See [19] for examples of formal languages with such expressive resources. Their discussion of what they call "external Curries," has some affinity to my discussion below, although it doesn't directly relate to metainferential phenomena.

[13]By "valid," I really mean here "derivable". For, as Hlobil argues in [10], a validity predicate cannot capture the admissible rules on pain of going non-classical in the metalanguage.

[14]This definition serves as an alternative to the explicit definition of validity in terms of truth preservation. To be more concrete, by "inferential definition" I mean an implicit definition of a specific kind: a concept is defined in such a way if it is given by rules that specify the concept's contribution to the premises and conclusions of inferences, i.e., its contribution to inferenes on both sides of the sequent arrow.

words, $V_0 P$ grants each validity statement the force of a theorem, equal in modal strength to the corresponding inference whose validity it's supposed to capture. On the other hand, $V_0 D$ is a detachment rule, whose rationale is analogous to that of *modus ponens*: it's nothing but plausible to infer a conclusion on the basis of certain premises together with the information that these premises entail that conclusion [6, pp. 152-153].

Although they have been contested (see, e.g., [3, pp. 10-15];[11, pp. 4765-4769]), $V_0 P$ and $V_0 D$ seem to maintain their appeal, due to the intuitive and straightforward definition they provide for the the concept of the validity of inferences. For that reason, I shall follow suit in defining the rest of the predicates, that are meant to capture "validity" at higher metainferential levels. I shall do so in two steps. First, I shall discuss in particular Val_1. Then it will become clear how to define the rest of the predicates by way of generalization.

Val_1 is supposed to capture the validity of metainferences of level 1. As such, its introduction rule—let us call it $V_1 P$—should allow us to express a Val_1 statement such as $\Rightarrow Val_1(\langle \Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k \rangle, \langle \Sigma \Rightarrow \Pi \rangle)$, and it should allow us to do so in cases (and only in cases) where the relevant metainference is derivable. In other words, the sequent $\Rightarrow Val_1(\langle \Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k \rangle, \langle \Sigma \Rightarrow \Pi \rangle)$ should be inferable from the metainference $\dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k}{\Sigma \Rightarrow \Pi}$. In our setting, the way to do that is by:

$$(2) \dfrac{(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k}{\Sigma \Rightarrow \Pi}}{(1) \dfrac{\emptyset}{\Rightarrow Val_1(\langle \Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k \rangle, \langle \Sigma \Rightarrow \Pi \rangle)}} \; V_1 P$$

In this way, $V_1 P$ turns out to be a mere generalization of $V_0 P$. In particular, observe that $V_0 P$ is $context-free$, i.e., it allows us to introduce validity-statements only without contexts, as the sole formula occurring in the sequent (with no additional premises or conclusions).[15] Analogously, in the case of metainferences, the "context" that has to be free is not only the sequent where Val_1 is introduced, but also the premise inference that precedes this sequent. The rationale is the same: each Val_1 statement has to be equal in modal strength with the metainference whose validity it's supposed to express.

As for $V_1 D$, the following detachment rule naturally suggests itself:

$$(1) \dfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k}{\Sigma, Val_1(\langle \Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_k \Rightarrow \Delta_k \rangle, \langle \Sigma \Rightarrow \Pi \rangle) \Rightarrow \Pi} \; V_1 D$$

That is to say, assuming that the premise sequents of a metainference hold, and that the premise of the conclusion sequent holds as well, together with the information that the metainference is valid, one is allowed to derive the conclusion of the conclusion sequent.

In general, we can follow the same line of thought and introduce sequent rules for Val_n for all $n \geq 1$. Let $\Gamma \Rightarrow_n A$ be some metainference of level n where

---

[15]For a justification of the context-freeness of $V_0 P$, see [3, p. 10].

$\Gamma = \{\gamma_1, ..., \gamma_m\}$, and let $\Sigma, \Pi \subseteq \mathscr{L}$ be some sets of formulas. Then, the following rule, $V_n P$, is a straightforward generalization of $V_1 P$ (where $\emptyset^n$ is the empty metainference of level n, namely, the metainference of level n whose premises and conclusions are all empty: $\emptyset^0 = \emptyset$, $\emptyset^1 = \frac{\emptyset}{\emptyset}$, $\emptyset^2 = \frac{\frac{\emptyset}{\emptyset}}{\frac{\emptyset}{\emptyset}}$, etc.):

$$(n+1) \quad \frac{\Gamma \Rightarrow_n A}{\emptyset^n (\Rightarrow Val_n(\langle\Gamma\rangle, \langle A\rangle))} \quad V_n P$$

As for $V_n D$, the following detachment rule naturally suggests itself:

$$(n) \quad \frac{\gamma_1, ..., \gamma_m}{A(\Sigma, Val_n(\langle\Gamma\rangle, \langle A(\Sigma \Rightarrow \Pi)\rangle)) \Rightarrow \Pi} \quad V_n D$$

That is to say, we can derive any metainference of level n of the form $\Gamma \Rightarrow_n A(\Sigma \Rightarrow \Pi)$, if we add the information that this metainference is valid (i.e., $Val_n(\langle\Gamma\rangle, \langle A(\Sigma \Rightarrow \Pi)\rangle)$) as a premise to the bottom sequent.

So much for definitions.[16] The question that arises, given those definitions, is whether one's logic is able to entertain them all, i.e., whether the *expressive power*[17] of one's logic allows one to define these validity predicates in that logic, capturing its own notions of "validity," regarding various metainferential levels. As is well known, regarding Val_0, the answer is not always positive. For, given self-reference,[18] there will be a sentence, κ_0, that is intersubstitutable with $Val_0(\langle\kappa_0\rangle, \langle\bot\rangle)$. Thus, one could reason as follows—let us call it derivation (0)—to derive the paradoxical sequent $\Rightarrow\bot$ (which, as is well known, leads to disaster):

$$\cfrac{\cfrac{\cfrac{\overline{\kappa_0, Val_0(\langle\kappa_0\rangle, \langle\bot\rangle) \Rightarrow\bot}}{\cfrac{\kappa_0 \Rightarrow\bot}{\Rightarrow Val_0(\langle\kappa_0\rangle, \langle\bot\rangle)}\,V_0 P}{\Rightarrow \kappa_0}\,\kappa_0\text{-}def \quad \cfrac{\overline{\kappa_0, Val_0(\langle\kappa_0\rangle, \langle\bot\rangle) \Rightarrow\bot}\,V_0 D}{\kappa_0 \Rightarrow\bot}\,\kappa_0\text{-}def}{\Rightarrow\bot}\,Cut$$

---

[16]Those definitions should be regarded as materially adequate, as they guarantee $V-schema$ for all levels. Here is a sketch of the proof: by $V_n D$ we get $\Gamma \Rightarrow_n A(\Sigma, Val_n(\langle\Gamma\rangle, \langle A(\Sigma \Rightarrow \Pi)\rangle)) \Rightarrow \Pi)$, and then

$$\cfrac{\Gamma \Rightarrow_n A(\Sigma, Val_n(\langle\Gamma\rangle, \langle A(\Sigma \Rightarrow \Pi)\rangle)) \Rightarrow \Pi) \quad \cfrac{\emptyset^n(\Rightarrow Val_n(\langle\Gamma\rangle, \langle A\rangle))}{\Gamma \Rightarrow_n A(\Sigma \Rightarrow Val_n(\langle\Gamma\rangle, \langle A(\Sigma \Rightarrow \Pi)\rangle), \Pi)}\,\substack{Enough\ applications\ of\ Weak_k\ (k \le n)}}{\Gamma \Rightarrow_n A(\Sigma \Rightarrow \Pi)}\,\substack{meta_n\text{-}Cut}$$

[17]By the "expressive power" of a logic, I mean here the set of operators that are "available" (using the terminology of [1, p. 119]) in that logic, where such "availability" is measured by the requirement that adding these operators to the language (along with their rules) doesn't trivialize the consequence relation at hand.

[18]By "self-reference," I mean expressive power rich enough to enable one to prove the diagonal lemma. That is to say, I assume that for any definable predicate in our language, P, there is a sentence A which "says of itself" that it has the property P, namely, $A \leftrightarrow P(\langle A\rangle)$ is a theorem. Moreover, I shall assume that A and $P(\langle A\rangle)$ are intersubstitutable on both sides of the turnstile. These assumptions are made for the sake of the argument, and for our purposes need not be justified here. See [8, pp. 853-855] for a discussion on how to achieve such results.

As a result, CL, as well as LP,[19] cannot be conservatively extended with Val_0, and are unable to capture their own notion of the validity of inferences. By contrast, the above derivation is blocked by ST, which is free of Cut. Indeed, ST can be conservatively extended with Val_0 just as it can be conservatively extended with a truth predicate [17, pp. 157-159].

But that's just the beginning of the story. For, given self-reference, there will be a sentence κ_1 that is intersubstitutable with $Val_1(\langle\emptyset\rangle, \langle\kappa_1 \Rightarrow\bot\rangle)$,[20] so that one can make derivation (1), which goes as follows:

$$
\cfrac{
  \cfrac{
    \cfrac{
      \cfrac{(1)\ \cfrac{\emptyset}{\kappa_1, Val_1(\langle\emptyset\rangle, \langle\kappa_1 \Rightarrow\bot\rangle) \Rightarrow\bot}\ V_1 D}{(2)\ \cfrac{}{\ }}\ \kappa_1-def
      }{(1)\ \cfrac{\emptyset}{\kappa_1 \Rightarrow\bot}}
    }{(2)\ \cfrac{(1)\ \cfrac{\emptyset}{\Rightarrow Val_1(\langle\emptyset\rangle, \langle\kappa_1 \Rightarrow\bot\rangle)}}{(1)\ \cfrac{\emptyset}{\Rightarrow \kappa_1}}\ \kappa_1-def}\ V_1 P
    \qquad
    \cfrac{(1)\ \cfrac{\emptyset}{\kappa_1, Val_1(\langle\emptyset\rangle, \langle\kappa_1 \Rightarrow\bot\rangle) \Rightarrow\bot}\ V_1 D}{(2)\ \cfrac{(1)\ \cfrac{\emptyset}{\kappa_1 \Rightarrow\bot}}{\ }}\ \kappa_1-def
  }{(1)\ \cfrac{\emptyset}{\Rightarrow\bot}}
}{}\ meta_1-Cut
$$

This is a disaster. For one thing, if the metalanguage is transitive, one can derive $\Rightarrow\bot$ from $\cfrac{\emptyset}{\Rightarrow\bot}$ by reasoning in the metalanguage. Worse, the conclusion metainference is not derivable in CL^1 and so, in a setting that is classical up to the first metainferential level—i.e., a setting where $meta_1 - Cut$ is available—we reach a contradiction. We thus get a version of the Curry paradox for metainferences of the first level.

Note that the above derivation is blocked in ST, because it depends on an application of $meta_1 - Cut$, which isn't available unless Cut is present (Proposition 3.3). The same is true of the third logic in the hierarchy L_2—also known as TS/ST—where $meta_1 - Cut$ fails to hold as well. Consequently, in the same way that ST blocks the well known Curry paradox, both ST and TS/ST block the aforesaid version of that paradox. For $n \geq 3$, however, $meta_1 - Cut$ is available, and so one may raise the conjecture that for $n \geq 3$, L_n cannot be conservatively extended with Val_1.

This result is naturally generalized in the following way. For all $n \geq 2$ we shall be able to make a corresponding derivation (n), for (given self-reference) there will be

a sentence κ_n that is intersubstitutable with $Val_n(\langle \emptyset^{n-1}\rangle, \langle \emptyset^{n-1}(\kappa_n \Rightarrow \bot)\rangle)$. This κ_n will give rise to the n-th version of the paradox:

$$
\cfrac{
\cfrac{
(n)\;\cfrac{\emptyset^{n-1}}{\emptyset^{n-1}(\kappa_n, Val_n(\langle \emptyset^{n-1}\rangle, \langle \emptyset^{n-1}(\kappa_n \Rightarrow \bot)\rangle)) \Rightarrow \bot}\; V_n D}{\emptyset^n(\kappa_n \Rightarrow \bot)}\;\kappa_n-def
}{
\cfrac{\emptyset^n(\Rightarrow Val_n(\langle \emptyset^{n-1}\rangle, \langle \emptyset^{n-1}(\kappa_n \Rightarrow \bot)\rangle))}{\emptyset^n(\Rightarrow \kappa_n)}\;\kappa_n-def
}\; V_n P
\qquad
(n)\;\cfrac{
\cfrac{\emptyset^{n-1}}{\emptyset^{n-1}(\kappa_n, Val_n(\langle \emptyset^{n-1}\rangle, \langle \emptyset^{n-1}(\kappa_n \Rightarrow \bot)\rangle)) \Rightarrow \bot}\; V_n D
}{\emptyset^n(\kappa_n \Rightarrow \bot)}\;\kappa_n-def
}{\emptyset^n(\Rightarrow \bot)}\; meta_n-Cut
$$

Consequently, it seems like each logic in the hierarchy can be conservatively extended with Val_n iff it doesn't admit the $meta_n - Cut$ rule. Truth be told, I was initially inclined to claim that:

Conjecture *For each $k \geq 1$, the logic L_k in the hierarchy can be conservatively extended with the validity predicate Val_n iff $k - 1 \leq n$.*

Yet, however tempting and elegant, at the end of the day this conjecture proves to be wrong. The fact of the matter is that each L_k (for $k \geq 1$) can conservatively extended with Val_n for all $n \in \mathbb{N} \cup \{0\}$. Let me explain why, taking TS/ST to be my case in point here. On the face of it, TS/ST has Cut and the rest of the rules that take part in derivation (0), so how can it block the paradox? The problem is that the proof tree of derivation (0) makes implicit use of transitivity rules. Each sequent in the tree, except for the leaves and the root, has a dual role to play: first, it is *inferred*, namely, it plays the role of a conclusion that results from applying some valid inference rules (given some premises). Then, it serves as a *premise* in yet another valid inference rule. However, TS/ST is non-transitive exactly in that it doesn't have the $tran_1$ rules, and so it doesn't allow us to make previously inferred sequents into premises of a new metainference. Suppose, for example, that we grant the derivations of $\Rightarrow \kappa_0, \kappa_0 \Rightarrow \bot$. With those sequents as premises, and with the following instance of Cut: $\cfrac{\Rightarrow \kappa_0, \kappa_0 \Rightarrow \bot}{\Rightarrow \bot}$, we want to derive $\Rightarrow \bot$. However, even though Cut holds in TS/ST, namely, $\vdash_{TS/ST} \cfrac{\Gamma, A \Rightarrow \Delta \quad \Gamma \Rightarrow A, \Delta}{\Gamma \Rightarrow \Delta}$, we are not allowed to apply it in the case of those premises because

$$
\nvdash_{TS/ST} \;(2)\; \cfrac{
(1)\;\cfrac{\emptyset}{\Gamma, A \Rightarrow \Delta}\quad (1)\;\cfrac{\emptyset}{\Gamma \Rightarrow A, \Delta}\quad (1)\;\cfrac{\Gamma, A \Rightarrow \Delta \quad \Gamma \Rightarrow A, \Delta}{\Gamma \Rightarrow \Delta}
}{(1)\;\cfrac{\emptyset}{\Gamma \Rightarrow \Delta}}
$$

As a result, TS/ST blocks derivation (0), not by disapproving some of the rules it involves, but by making it impossible to concatenate these rules into one proof tree. Likewise, although L_3 does approve the above principle, it doesn't allow us to apply it to guarantee that derivation (0) goes through. In general, as Scambler stresses in [18], each L_n (for $n \geq 2$) accepts Cut, but denies metaprinciples that allow us to go from the acceptance of Cut and the acceptance of some Cut premises to the acceptance of the relevant Cut conclusion. For this (very interesting) reason, derivation (0) is

blocked by each such L_n (for $n \geq 2$): these logics explicitly seem to recognize that the curry sentence leads to disaster, while also *not* allowing it to lead to disaster. Likewise, every other derivation (n) is blocked by all the logics (for all $n \in \mathbb{N} \cup \{0\}$): up to level $n-1$ by not having $meta_n - Cut$, and from there onwards, by denying structural principles that will allow us to use $meta_n - Cut$ in the derivation. Thus, it turns out that each logic in the hierarchy (except for LP) can handle all the above validity predicates.[21]

5 Philosophical Implications

In this section, I wish to point out two philosophical implications of the above results. I don't pretend to be arguing these implications conclusively. Rather, my goal is merely to demonstrate the philosophical potential there is in the proof-theoretic account of the BPS hierarchy. A more comprehensive discussion of these implications will have to wait for another occasion.

5.1 Identifying Logics

I take it that the proof-theoretic account makes it really hard to identify ST with CL, LP, or any other logic in the hierarchy. My reasons are as follows.

A priori, there are two ways (as well as combinations thereof) in which the identity of a logic may be determined: either *extensionally*, in terms of some relevant extension, e.g., the extension of the logic's consequence relation (which sets $\Gamma, \Delta \subseteq \mathscr{L}$ stand in the $\Rightarrow$ relation), or *intensionally*, in terms of the logic's rules that determine such an extension. Those who identify ST with CL clearly reside in the second camp. For them, although ST and CL differ on the rules—and thereby differ also in expressive power—they have the same consequence relation, and that is all that matters. A difference of opinion regarding the rules—particularly, as to whether Cut should be derivable—doesn't really count on this view unless it affects the consequence relation at hand.

Now, although the nature of the metainferential phenomena doesn't speak against the extensional criterion as such, it does speak against the claim that the relevant extension—that which determines a logic's identity by this criterion—is just that of the consequence relation. For, if the hierarchy has taught us anything, it is that a comprehensive account of logic has to include sequent rules and axioms governing not only inferences, but metainferences of all levels, so as to determine not only the consequence relation of the logic at hand, but also the extension of the logic's valid metainferences, metametainferences, etc. Namely, all metainferential levels should

---

[21]The informed reader may be reminded at this point of a proof given by Pailos in [13], that the same goes for a transparent truth predicate. Namely, each such logic can be conservatively extended with a predicate T such that $v(T(\langle A \rangle)) = v(A)$ for all valuations, in all models. See also the discussion in [18], particularly Section 5 there.

be taken into account in specifying the extension that determines a logic's identity.[22] Therefore, ST is no longer identifiable with CL because these logics differ extensionally at infinitely many metainferential levels, nor are any two logics in the hierarchy identifiable for the same reason.

Therefore, it no longer seems viable to identify ST with CL, either extensionally or intensionally, nor does it seem viable to identify any two logics in the hierarchy with one another, either extensionally, or intensionally, because all these logics have at least some different rules. Thus, it seems that each logic maintains its own distinct identity.

How about identifying some of these logics through translation functions? The idea behind such an identification is that two logics are identifiable if they have the same "abstract consequence relation," as defined, e.g., by Blok and Jónsson in [7]. However, from a philosophical point of view, it is quite clear that the identification of two logics (or any other pair of objects, for that matter) through translation functions already inclines us toward the intensional criterion: it makes no sense to identify two objects as the same through a translation function if the identity of these objects is determined extensionally and the two relevant extensions are (set-theoretically) different. That is to say, the very appeal to translation functions in order to determine the identity of a given object concedes that such an identity is determined in terms of *rules*, rather than in terms of extensions. It follows that the relevant translation function to which we appeal must not only preserve some extension; the rules that determine this extension must also be somehow preserved. Yet it is unclear in what sense a translation function can "preserve" inference rules.[23] Thus, despite being of *mathematical* interest, the translation function by itself falls short of establishing a *metaphysical* claim of identity. Since LP and ST (in fact, all the logics in the hierarchy) differ on the rules, it seems that translation functions are of little help in establishing claims about their identity.[24]

---

[22]Note that the rules that determine the extension at level n (which are schemata of metainferences of level $n+1$) are themselves regarded at level $n+1$ as axioms, namely, as part of the "extension". Thus, we face here the phenomenon already observed by Lewis in [12], that "intensional" is a relative concept, and cannot be simply defined as "that which is not extensional".

[23]See [20, pp 63-64] for a discussion of relevant problems. Indeed, according to Woods' criterion of intertranslatability, ST and LP aren't identical.

[24]The situation is even worse in the case of ST and LP, which clearly differ in expressive power, since only the former can handle a validity predicate. Moreover, the claim that a logic is to be identified with its external logic faces yet another challenge here. For, the external logic of each logic in the hierarchy is its predecessor. Hence, if a system is to be identified with its external logic, we would have to identify not only ST with LP, but also L_m with L_{m-1} for all $m \in \mathbb{N}$. Then, provided that "identity" is a transitive relation, we end up in the position where all of these logics are identical to LP, which may strike one as odd. Furthermore, it becomes really hard to argue that LP is different from CL, since for each level of "closeness" to CL there is a logic in the hierarchy that is identical to LP. To put the matter topologically (which seems appropriate here), LP has a representative in any "neighborhood" of CL and so it is as "close" to CL as possible. This conclusion is at least odd, if not unacceptable.

5.2 Each Validity Predicate is Required

In [2], Barrio, Rosenblatt, and Tajer put forward a criticism of ST, targeting its alleged ability to handle a validity predicate. First, they point out that the presence of a validity predicate in the object language makes it possible to express not only validity statements such as $Val_0(\langle A \rangle, \langle A \vee B \rangle)$, but also the metarules, e.g., $L\neg$: $Val_0(\langle \Gamma \rangle, \langle \Delta \cup \{A\} \rangle) \Rightarrow Val_0(\langle \Gamma \cup \{\neg A\} \rangle, \langle \Delta \rangle)$. That is, ST enriched with a validity predicate seems to "have an opinion" on whether such metarules are valid.

However, the criticism goes, it is impossible to derive sequents like $Val_0(\langle \Gamma \rangle, \langle \Delta \cup \{A\} \rangle) \Rightarrow Val_0(\langle \Gamma \cup \{\neg A\} \rangle, \langle \Delta \rangle)$ in ST. This is because ST is Cut-free, and so the only way to introduce such a Val_0 statement must come by $V_0 P$ and $V_0 D$. Yet, $V_0 P$ is context-free, which makes it impossible to derive such a sequent with it (see the details in [3, p. 10 (Proposition 1)]). Alternatives to $V_0 P$ prove even more problematic, as they allow us to derive an internalized version of Cut [3, pp. 15-16]. Thus, it seems that ST cannot properly handle a validity predicate because any such predicate fails to "have the right opinion" on ST's metarules.

With the hierarchy, the situation is even worse. For, if it is impossible to derive statements like $Val_0(\langle \Gamma \rangle, \langle \Delta \cup \{A\} \rangle) \Rightarrow Val_0(\langle \Gamma \cup \{\neg A\} \rangle, \langle \Delta \rangle)$, which states that $L\neg$ is valid, it is definitely impossible—for the same reason, the context-freeness of $V_0 P$—to derive statements like:

$$Val_0(\langle \Gamma \cup \{\neg A \vee \neg B\} \rangle, \langle \Delta \rangle) \Rightarrow Val_0(\langle \Gamma \cup \{\neg (A \wedge B)\} \rangle, \langle \Delta \rangle)$$

which says that the de-Morgan rule $\neg \wedge Left$ is valid. For, the derivation of $\neg \wedge Left$ in ST involves $L\neg$. As a result, it is also impossible to derive the following sequent: $Val_0(\langle\langle (Val_0(\langle \emptyset \rangle, \langle \gamma_1 \rangle) \wedge ... \wedge Val_0(\langle \emptyset \rangle, \langle \gamma_m \rangle)) \wedge Val_0(\langle \emptyset \rangle, \langle \neg A \vee \neg B \rangle) \rangle, \langle Val_0(\langle \emptyset \rangle,$ $\langle \Delta \rangle) \rangle) \Rightarrow Val_0(\langle\langle (Val_0(\langle \emptyset \rangle, \langle \gamma_1 \rangle) \wedge ... \wedge Val_0(\langle \emptyset \rangle, \langle \gamma_m \rangle)) \wedge Val_0(\langle \emptyset \rangle, \langle \neg (A \wedge B) \rangle) \rangle,$ $\langle Val_0(\langle \emptyset \rangle, \langle \Delta \rangle) \rangle)$ (where $\Gamma = \{\gamma_1, ..., \gamma_m\}$), which is supposed to state that $higher(\neg \wedge Left)$ is derivable in ST^∞. But $higher(\neg \wedge Left)$ is derivable in ST^∞ (Theorem 3.10). Thus, Val_0 fails to capture not only the metarules, but derivable metainferences of levels> 1 as well.[25]

By contrast, Val_1 is perfectly suitable for capturing the validity of metarules (of level 1) such as $L\neg$; it simply does so by $Val_1(\langle \Gamma \Rightarrow A, \Delta \rangle, \langle \Gamma, \neg A \Rightarrow \Delta \rangle)$ and $V_1 P$:

$$(2) \cfrac{(1) \cfrac{\cfrac{\Gamma \Rightarrow A, \Delta}{\Gamma, \neg A \Rightarrow \Delta} \; L\neg}{\emptyset}}{\Rightarrow Val_1(\langle \Gamma \Rightarrow A, \Delta \rangle, \langle \Gamma, \neg A \Rightarrow \Delta \rangle)} \; V_1 P$$

In fact, Val_1 captures in such a way all valid metainferences of level 1. Likewise, Val_2 is suitable for capturing the validity of $higher(\neg \wedge Left)$ and, in general, ST^∞ is capable of expressing the validity of all of its higher-order metainferences, in each case with the appropriate validity predicate. In sum, the criticism of ST rests on an unjustified assumption: it's not that ST fails to have the right opinion on its

---

[25] This difficulty applies also to alternative ways to deal with Barrio, Rosenblatt, and Tajer's criticism. See, e.g., Hlobil's suggestion in [11], in particular Section 5 there.

own metarules; it's just that expressing such an opinion requires yet another validity predicate, not the one used to express what follows from what.

6 Conclusion

The purpose of this paper was to explore the BPS hierarchy from a proof-theoretic perspective. This perspective is important for at least three reasons. First, the proof-theoretic account is complementary to the model-theoretic account. Hence, this account puts us in a position to gain a better understanding of the hierarchy, both technically and conceptually. Second, the proof-theoretic account allows us to define the sequence of validity predicates, and to use it to study the expressive power of the logics in the hierarchy. Third, the proof-theoretic account allows us to draw some philosophical conclusions about the hierarchy. In particular, I am inclined to think that each logic in the hierarchy maintains its own distinct identity, and that a validity predicate need not capture "validity" at more than one metainferential level.

Appendix

Lemma 2.8. *For all $n \geq 1$, and for all $\Gamma \subseteq SEQ^{n-1}(\mathscr{L})$, $A \in SEQ^{n-1}(\mathscr{L})$:*
$\models_{CL} \Gamma \Rightarrow_n A$ iff $\models_{CL} lower(\Gamma \Rightarrow_n A)$.

Proof By Theorem 2.6 $\models_{CL} \Gamma \Rightarrow_n A$ iff $\models_{L_{n+1}} \Gamma \Rightarrow_n A$. By Theorem 2.7 the latter is the case iff $\models_{L_n} lower(\Gamma \Rightarrow_n A)$, and once again by Theorem 2.6 (since $lower(\Gamma \Rightarrow_n A)$ is a metainference of level $n-1$), that's the case iff $\models_{CL} lower(\Gamma \Rightarrow_n A)$. $\qquad\qquad\square$

Proposition 3.3. *Here are some useful facts about these axioms:*

(i) *$Weak_n^b$ for $n \geq 1$ is redundant: given the well-known Weakening rule (for inferences), it follows by induction for all n, via $Tran_n^c$.*

(ii) *For each instance of such axioms/rules r of level n (except for $Weak_n^b$ which is redundant, given (i)), $lower(r)$, $higher(r)$ instantiate that same axiom/rule at levels $n-1$, $n+1$, respectively.*

(iii) *Given Cut, $meta_n$-Cut (for $n \geq 1$) is derivable via $Tran_n^c$.*

Proof (i) By induction. Assuming $Weak_n^b$ is derivable, we can derive $Weak_{n+1}^b$ from $Weak_n^b$ and $Tran_{n+1}^c$ in the following way:

$$(n+2) \cfrac{(n+1)\cfrac{\Gamma_1 \Rightarrow_n A_1, ..., \Gamma_k \Rightarrow_n A_k}{\Sigma \Rightarrow_n B(\Lambda \Rightarrow \Theta)} \quad (n+1)\cfrac{\Sigma \Rightarrow_n B(\Lambda \Rightarrow \Theta)}{\Sigma \Rightarrow_n B(\Lambda, \Lambda' \Rightarrow \Theta, \Theta')}\ Weak_n^b}{(n+1)\cfrac{\Gamma_1 \Rightarrow_n A_1, ..., \Gamma_k \Rightarrow_n A_k}{\Sigma \Rightarrow_n B(\Lambda, \Lambda' \Rightarrow \Theta, \Theta')}}\ Tran_{n+1}^c$$

Part (ii) is easy to prove, and so it is left for the reader as an exercise.

Part (iii) is proven by induction on n. The base case, Cut, is given. Assume that the claim holds for n, and let $\Gamma = \{\gamma_1, ..., \gamma_m\} \subseteq SEQ^n(\mathscr{L})$, $A \in SEQ^n(\mathscr{L})$. Then $meta_{n+1} - Cut$ is derivable from $meta_n - Cut$ (guaranteed by the inductive hypothesis) and $Tran_{n+1}^c$:

$$(n+1)\cfrac{(n+1)\cfrac{\gamma_1, ..., \gamma_m}{A(\Sigma, B \Rightarrow \Pi)} \quad (n+1)\cfrac{\gamma_1, ..., \gamma_m}{A(\Sigma \Rightarrow B, \Pi)} \quad (n+1)\cfrac{A(\Sigma, B \Rightarrow \Pi) \quad A(\Sigma \Rightarrow B, \Pi)}{A(\Sigma \Rightarrow \Pi)}\small{Tran_{n+1}^c} \normalsize{}_{meta_n - Cut}}{\gamma_1, ..., \gamma_m \over A(\Sigma \Rightarrow \Pi)}$$

$\square$

Proposition 3.5. *For all $n \in \mathbb{N} \cup \{0\}$, and for every metainference $\Gamma \Rightarrow_n A$: $\vdash_{CL^\infty} \Gamma \Rightarrow_n A$ iff $\vdash_{CL^\infty} lower(\Gamma \Rightarrow_n A)$.*

Proof By induction. I shall prove the claim for $n = 0, 1, 2$, and then turn to the inductive step.

The $n = 0$ case is the deduction theorem, which is known to hold in CL, so $\vdash_{CL^\infty} \Gamma \Rightarrow \Delta$ iff $\vdash_{CL^\infty} (\wedge\Gamma) \supset (\vee\Delta)$, which is $\vdash_{CL^\infty} lower(\Gamma \Rightarrow \Delta)$.

The $n = 1$ case follows from the deduction theorem along with the fact that CL coincides with its external logic. Let $\cfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi}$ be some metainference of level 1. By the deduction theorem, for all $1 \leq i \leq m$: $\vdash_{CL^\infty} \Gamma_i \Rightarrow \Delta_i$ iff $\vdash_{CL^\infty} lower(\Gamma_i \Rightarrow \Delta_i)$, and $\vdash_{CL^\infty} \Sigma \Rightarrow \Pi$ iff $\vdash_{CL^\infty} lower(\Sigma \Rightarrow \Pi)$. Now, by $m + 1$ applications of the transitivity rules, we can replace each of these sequents with its equivalent one, and get that

$$\vdash_{CL^\infty} \cfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi} \text{ iff } \vdash_{CL^\infty} \cfrac{\Rightarrow lower(\Gamma_1 \Rightarrow \Delta_1), ..., \Rightarrow lower(\Gamma_m \Rightarrow \Delta_m)}{\Rightarrow lower(\Sigma \Rightarrow \Pi)}.$$

Since CL coincides with its external logic, the latter holds iff

$$\vdash_{CL^\infty} lower(\Gamma_1 \Rightarrow \Delta_1), ..., lower(\Gamma_m \Rightarrow \Delta_m) \Rightarrow lower(\Sigma \Rightarrow \Pi)$$

which is $\vdash_{CL^\infty} lower(\cfrac{\Gamma_1 \Rightarrow \Delta_1, ..., \Gamma_m \Rightarrow \Delta_m}{\Sigma \Rightarrow \Pi})$.

We also need to prove the $n = 2$ case. It is enough to show that for each instance r of a rule/axiom of level 2, $lower(r)$ is derivable in CL^∞, and that for each instance q of a rule/axiom of level 1, $higher(q)$ is derivable in CL^∞. If so, then:

1. For each proof tree of metainferences of level 2 whose leaves are $r_1, ..., r_n$ and whose root is b there is a proof tree of metainferences of level 1 (a "regu-

lar" proof tree) whose leaves are $lower(r_1), ..., lower(r_n)$ and whose root is $lower(b)$.

2. For each proof tree of metainferences of level 1 whose leaves are $l_1, ...l_m$ and whose root is q, there is a proof tree of level 2 whose leaves are $higher(l_1), ..., higher(l_m)$ and whose root is $higher(q)$.

Let us begin with the first claim. Consider an instance r of some rules/axiom of level 2. It is necessarily one of the structural rules, and so by Proposition 3.3 $lower(r)$ also instantiates that same axiom/rule at level 1.

Consider then the axioms and rules of level 1, namely, the regular sequent rules and axioms of CL. We have to show that for each instance of some such axiom/rule q, $higher(q)$ is derivable in CL^∞. I shall discuss two cases, and leave the rest as an exercise.

$$(i)\ higher(R\vee) = {}^{(2)}\dfrac{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A(B), \Delta}}{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A \vee B, \Delta}}\quad \text{can be derived in the following}$$

way:[26]

$$^{(2)}\dfrac{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A(B), \Delta}\quad {}^{(1)}\dfrac{\Rightarrow A(B), \Delta}{\Rightarrow A \vee B, \Delta}\ {}_{R\vee}}{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A \vee B, \Delta}}\ {}_{Tran_1^c} \qquad {}^{(2)}\dfrac{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A \vee B, \Delta}\quad {}^{(1)}\dfrac{\Rightarrow A \vee B, \Delta}{\Rightarrow A(B), \Delta}\ {}_{R\vee}}{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A(B), \Delta}}\ {}_{Tran_1^c}$$

(ii) The proof of the $higher(R\neg)$ case is more complex. I shall discuss the top-to-bottom direction, but the proof of the other direction is analogous. What we need to prove is that for all Γ, Δ, A:

$$^{(2)}\dfrac{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m, \Rightarrow A}{\Rightarrow \Delta}}{{}^{(1)}\dfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow \neg A, \Delta}}$$

---

[26]Recall that $R\vee$ is a double-line rule in the calculus given in Section 2.

What we really need to show is that the metarule *push*:

$$(2) \; \cfrac{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m, \Rightarrow A}{\Rightarrow \Delta}}{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{A \Rightarrow \Delta}} \quad push \text{ is admissible in } CL^\infty.^{27} \text{ If so, then}$$

whatever is derivable with *push* will be derivable without it. And so, for all Γ, Δ, A:

$$(2) \; \cfrac{(2) \; \cfrac{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m, \Rightarrow A}{\Rightarrow \Delta}}{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{A \Rightarrow \Delta}} \; push \qquad (1) \; \cfrac{A \Rightarrow \Delta}{\Rightarrow \neg A, \Delta} \; R\neg}{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow \neg A, \Delta}} \; Tran_1^c$$

But actually, *push* is admissible simply because CL coincides with its external logic.[28] That completes the proof of the $n = 2$ case.

Now, suppose that the claim holds for some $n \geq 2$. We want to show that it holds for $n + 1$. Since all the rules and axioms of levels $n, n + 1$ are structural, this is an

---

[27] Admissibility is defined for metainferences in the same way that it is defined for inferences: an admissible rule is one whose addition to the existing rules of the system doesn't change the system's set of theorems, or, in this case, the set of derivable metainferences. Thus, to prove that a rule is admissible, we need to show that if certain instances of the premise metainferences hold, the relevant instance of the conclusion metainference holds as well.

[28] Nonetheless, here is a sketch of a proof that *push* is admissible: Assume that the (meta)sequent $\cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m, \Rightarrow A}{\Rightarrow \Delta}$ holds. So there is a proof tree whose leaves are $\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m, \Rightarrow A$ and whose root is $\Rightarrow \Delta$. We show by induction on the proof height how to construct a proof tree of $\cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{A \Rightarrow \Delta}$. For a height of 1, if the sequent $\Rightarrow A$ isn't active in the proof, then there is a proof tree of $\cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow \Delta}$. Using $Weak_1^b$, we can weaken with A on the left of the bottom sequent and get the result. If $\Rightarrow A$ is active in the proof, then there are three options: Δ is derived either by Ref_1 or by $\vee R$ or by $\wedge R$. In the first case $A \in \Delta$ and so $A \Rightarrow \Delta$ is an identity axiom (and from there the proof is trivial). In the second case, there is some B such that $A \vee B \in \Delta$. So $\cfrac{A \Rightarrow A}{A \Rightarrow (A \vee B)\Delta} \; \vee R$, and from there with the $Weak_1$ rules the proof is trivial.

In the third case, there are two options: (i) there is some $1 \leq i \leq m$ such that $\gamma_i \wedge A \in \Delta$, and (ii) A is of the form $B \wedge C$ and B (*or* C) $\in \Delta$. In the case of (i) we reason:

$$\cfrac{A \Rightarrow A \quad \cfrac{\Rightarrow \gamma_i}{A \Rightarrow \gamma_i} \; Weakening}{A \Rightarrow (A \wedge \gamma_i)\Delta} \; \wedge R, \, Weakening$$

And once again, enough applications of the $Weak_1$ rules will get the job done. As for (ii), $B \wedge C \Rightarrow B(C)\Delta$ is clearly derivable, and then we weaken m times with $\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m$ respectively. The inductive step is analogous for, given Cut-elimination, we can assume that the proof involves only sequents with an empty left-hand side.

immediate result of Proposition 3.3: for each instance r of some axiom/rule of level n, $higher(r)$ instantiates that same axiom/rule at level $n + 1$, and for each instance q of some axiom/rule of level $n + 1$, $lower(q)$ instantiates that same axiom/rule at level n. This completes the entire proof. $\qquad\square$

Theorem 3.11. *For all $m \geq 2$, for all $n \in \mathbb{N} \cup \{0\}$, and for every metainference $\Gamma \Rightarrow_n A$: $\vdash_{L_m^\infty} \Gamma \Rightarrow_n A$ iff $\vdash_{L_{m-1}^\infty} lower(\Gamma \Rightarrow_n A)$.*

Proof By nested induction.

Base case: $m = 2, n = 0$. The relevant rules governing inference (metainference of level 0) are those of CL^0. But it is known that $\vdash_{CL} \Gamma \Rightarrow \Delta$ iff $\vdash_{ST} lower(\Gamma \Rightarrow \Delta)$. I shall also discuss the $n = 1$ case. Consider a proof tree of a metainference (of the first level) $\Gamma \Rightarrow_1 A$. Now, by Proposition 3.3: $L_2^1 = CL^0 \cup Weak_1 \cup Ref_1 \cup Aux_1 \cup Sub_1$ and so for each instance of axiom/rule $r \in L_2^1$: $\vdash_{ST^0} lower(r)$. Hence, $\vdash_{ST^0} lower(\Gamma \Rightarrow_1 A)$. We also need to show that for each instance of axiom/rule $q \in ST^0$, $higher(q)$ is derivable in L_2^1. Once again, the proof is similar to that of Proposition 3.5: it relies on the fact that all the operational rules of ST^0, which are also the operational rules of L_2^1, are double-line rules. Therefore, it is easy to prove that their $higher$ versions are derivable, using instances of Sub_1. For example, both directions of $higher(R\vee)$ are derivable as follows:

$$
2) \; \cfrac{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A(B), \Delta} \quad (1) \; \cfrac{\Rightarrow A(B)}{\Rightarrow A \vee B)} \, R\vee}{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A \vee B, \Delta}} \, Sub_1^c
\qquad
(2) \; \cfrac{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A \vee B, \Delta} \quad (1) \; \cfrac{\Rightarrow A \vee B}{\Rightarrow A(B)} \, R\vee}{(1) \; \cfrac{\Rightarrow \gamma_1, ..., \Rightarrow \gamma_m}{\Rightarrow A(B), \Delta}} \, Sub_1^c
$$

Hence, $\vdash_{L_2^1} \Gamma \Rightarrow_1 A$ iff $\vdash_{ST^0} lower(\Gamma \Rightarrow_1 A)$.

For $n > 1$, however, observe that L_2^n and ST^{n-1} have only structural rules at these levels, and so the proof is trivial, given Proposition 3.3.

Inductive step: Assume that the claim holds for some m; then we have to show that it holds for $m + 1$, for all n. For $0 \leq n \leq m - 1$, the relevant rules are those of CL^{m-1}, and we know from Proposition 3.5 that $\vdash_{CL^{m-1}} \Gamma \Rightarrow_n A$ iff $\vdash_{CL^{m-2}} lower(\Gamma \Rightarrow_n A)$. For $n = m$, note that L_{m+1}^m and L_m^{m-1} have the same rules (modulo degrees) which are all structural (at level m of L_{m+1}^m and at level $m - 1$ of L_m^{m-1}), and so, once again, the claim holds, by Proposition 3.3. The same method of proof also applies to all $n > m$. This completes the proof. $\qquad\square$

References

1. Avron, A. (1991). Simple consequence relations. *Information and Computation*, *92*, 105–139.
2. Barrio, E., Rosenblatt, L., & Tajer, D. (2015). The logics of strict-tolerant logic. *Journal of Philosophical Logic*, *44*(5), 551–571.
3. Barrio, E., Rosenblatt, L., & Tajer, D. (2016). Capturing naive validity in the Cut-free approach. Synthese. https://doi.org/10.1007/s11229-016-1199-5.
4. Barrio, E., Pailos, F., & Szmuc, D. (2020). A hierarchy of classical and paraconsistent logics. *Journal of Philosophical Logic*, *49*, 93–120. https://doi.org/10.1007/s10992-019-09513-z.
5. Beall, J. (2011). Multiple-conclusion LP and default classicality. *Review of Symbolic Logic*, *4*(2), 326–336. https://doi.org/10.1017/S1755020311000074.
6. Beall, J., & Murzi, J. (2013). Two flavors of Curry paradox. *Journal of Philosophy*, *110*(3), 143–165.
7. Blok, W., & Jónsson, B. (2006). Equivalence of consequence operations. *Studia Logica*, *83*(1), 91–110.
8. Cobreros, P., Égré, P., Ripley, D., & van Rooij, R. (2013). Reaching transparent truth. *Mind*, *122*(488), 841–866.
9. Dicher, B., & Paoli, F. (2019). ST, LP, and tolerant metainferences. In Başkent, C., & Ferguson, T.M. (Eds.) Graham Priest on Dialetheism and Para consistency. Dordrecht: Springer, pp.383–407.
10. Hlobil, U. (2018). The Cut-free approach and the admissibility-Curry. Thought: a Journal of Philosophy, 7: 40–48. https://doi.org/10.1002/tht3.267.
11. Hlobil, U. (2019). Faithfulness for Naive Validity. *Synthese*, *196*, 4759–4774.
12. Lewis, D. (1974). Tensions. In Munitz, M.K. & Unger, P.K. (Eds.), *Semantics and Philosophy*. New York: New York University Press, pp. 49–61.
13. Pailos, F. (2020). A fully classical truth theory characterized by substructural means. *Review of Symbolic Logic*, *13*(2), 249–268.
14. Priest, G. (1979). The logic of paradox. *Journal of Philosophical Logic*, *8*, 219–241.
15. Priest, G. (2008). An Introduction to Non-Classical Logic, second edition. Cambridge, MA: Cambridge University Press.
16. Ripley, D. (2012). Conservatively extending classical logic with transparent truth. *Review of Symbolic Logic*, *5*(2), 354–378.
17. Ripley, D. (2013). Paradoxes and Failures of Cut. *Australasian Journal of Philosophy*, *91*(1), 139–164.
18. Scambler, C. (2020). Classical logic and the strict tolerant hierarchy. *Journal of Philosophical Logic*, *49*, 351–370.
19. Wansing, H., & Priest, G. (2015). External Curries. *Journal of Philosophical Logic*, *44*(4), 453–471.
20. Woods, J. (2018). Intertranslatability, theoretical equivalence, and perversion. *Thought: a Journal of Philosophy*, *7*, 58–68.