

# An Analysis of Gödel's Dialectica Interpretation via Linear Logic

Paulo Oliva

Department of Computer Science  
Queen Mary, University of London  
London E1 4NS, UK

pbo@dcs.qmul.ac.uk  
<http://www.dcs.qmul.ac.uk/~pbo>

**Abstract.** This article presents an analysis of Gödel's Dialectica interpretation via a refinement of intuitionistic logic known as linear logic. Linear logic comes naturally into the picture once one observes that the structural rule of contraction is the main cause of the lack of symmetry in Gödel's interpretation. We use the fact that the Dialectica interpretation of intuitionistic logic can be viewed as a composition of Girard's embedding of intuitionistic logic into linear logic followed by de Paiva's Dialectica interpretation of linear logic. We then investigate the various properties of the Dialectica interpretation, such as the characterisation theorem, and variants of Gödel's interpretation within the linear logic context. The role of contraction in extensions to classical logic, arithmetic and analysis is also discussed.

## 1 Introduction

This article aims at analysing Gödel's Dialectica interpretation [14] via a refinement of intuitionistic logic known as linear logic [11, 12]. More precisely, we discuss how Gödel's ingenious interpretation can be obtained via a combination of de Paiva's intuitive interpretation [25, 26] of linear logic and Girard's embedding [11] of intuitionistic logic into linear logic. By breaking the Dialectica interpretation into two well-defined steps we are able to give an analysis of the characterisation principles required by Gödel's interpretation and to gain flexibility in the extensions of the interpretation to arithmetic and analysis. We also investigate some of the variants of the Dialectica interpretation within the linear logic context. Our analysis is based on recent work of Shirahata [30]. Our main contribution with respect to de Paiva's and Shirahata's work is the characterisation theorem of the linear logic Dialectica interpretation, and the uniform treatment of the several variants of the Dialectica interpretation [22, 24].

Gödel's Dialectica interpretation of a formal system  $\mathcal{S}$  into another system  $\mathcal{T}$  associates each formula  $A \in \mathcal{L}(\mathcal{S})$  with a decidable binary relation  $A_D \subseteq \rho \times \tau$  in  $\mathcal{L}(\mathcal{T})$ . By  $\rho$  and  $\tau$  we mean sequences of finite types. The system  $\mathcal{S}$  is normally called the *interpreted system* while  $\mathcal{T}$  is called the *verifying system*. Intuitively, the binary relation  $A_D$  describes the adjudication relation in a one-move game between two players (say Eloise and Abelard) whose moves are taken from the finite types  $\rho$  and  $\tau$  respectively.

	$\mathcal{S}$	$\hookrightarrow$	$\mathcal{T}$
Formulas as games	$A$		$A_D(\mathbf{x}; \mathbf{y})$
Proofs as winning moves	$\vdash A$		$\vdash \forall \mathbf{y} A_D(\mathbf{a}; \mathbf{y})$

**Table 1.** The Dialectica interpretation

The game goes as follows: Eloise chooses an element  $\mathbf{a} \in \rho$ , Abelard then chooses an element  $\mathbf{b} \in \tau$ , Eloise wins if  $A_D(\mathbf{a}; \mathbf{b})$  holds, otherwise Abelard wins. The interpretation is such that from a proof of  $A$  in  $\mathcal{S}$  one can extract a winning move  $\mathbf{a} \in \rho$  for Eloise together with a proof that  $\mathbf{a}$  is a winning move, i.e.  $\forall \mathbf{y} A_D(\mathbf{a}; \mathbf{y})$ , in the verifying system. More precisely, a formula  $A$  is interpreted as the existence of a winning move for Eloise in the game  $A_D(\mathbf{x}; \mathbf{y})$ , i.e.

$$\exists \mathbf{x} \forall \mathbf{y} A_D(\mathbf{x}; \mathbf{y}) \quad (1)$$

and a proof of  $A$  provides a concrete winning move for Eloise. Notice that the tuples  $\mathbf{x}$  and  $\mathbf{y}$  could possibly be empty, in which case the corresponding player is not asked to make a move, but will win in case the truth value of the adjudication relation goes in her/his way. For instance, consider the trivial case of the formula  $A \equiv \forall x(x = x)$ . The formula  $A$  is associated with the game in which Eloise does not need to make a move, while Abelard tries to find a value  $b$  which refutes  $\forall x(x = x)$ , i.e. such that  $b \neq b$ . Since Abelard will never be able to find such  $b$ , Eloise wins the corresponding game no matter what choice Abelard makes.

In Gödel's seminal work,  $\mathcal{S}$  was first-order intuitionistic arithmetic and  $\mathcal{T}$  an extension of primitive recursive arithmetic to all finite types. Gödel's goal for developing the interpretation was to provide a relative consistence proof for first-order intuitionistic arithmetic. That is achieved since falsity  $\perp$  is interpreted as the game where none of the players need to make a move, and Abelard always wins. But, recall that a proof of  $A$  in  $\mathcal{S}$  gives rise to a winning move for Eloise and a verification of this fact in  $\mathcal{T}$ . Assuming  $\mathcal{T}$  is consistent Eloise cannot win the game corresponding to  $\perp$ , which implies that  $\perp$  cannot be derived in the interpreted system  $\mathcal{S}$ , i.e.  $\mathcal{S}$  is consistent.

There are, however, several issues which make the Dialectica interpretation difficult to understand or justify. First, there is the strange lack of symmetry between Eloise and Abelard, since Eloise is always required to make the first move and Abelard's move can depend on Eloise's move. This seems to make Abelard's life easier, and consequently, Eloise's task more difficult. The asymmetry in the roles of the players comes from the asymmetry in intuitionistic logic between assumptions and conclusions. Namely, in intuitionistic logic one tries to derive a *single* formula  $A$  from a *finite set* of assumptions  $\Gamma$ . In particular, a single assumption  $B$  can be used repeatedly in a proof in different ways in order to derive the single conclusion  $A$ . In other words, in intuitionistic logic one allows the structural rule of contraction for formulas in the premise

$$\frac{\Gamma, B, B \vdash A}{\Gamma, B \vdash A} \quad (2)$$

but not in the conclusion, by restricting the conclusion to contain a single formula.

Another issue with the Dialectica interpretation concerns the interpretation of implications  $A \rightarrow B$ . Given that interpretations for  $A$  and  $B$  in the form (1) have already been obtained, the formula

$$\exists x \forall y A_D(x; y) \rightarrow \exists v \forall w B_D(v; w) \quad (3)$$

is interpreted as

$$\exists f, g \forall x, w (A_D(x; fwx) \rightarrow B_D(gx; w)). \quad (4)$$

Despite empirical evidence that this is a choice of interpretation with excellent properties (cf. Bishop [5] and Kreisel [19]), it is difficult to argue why that should be so. All we can say is that in order to go from (3) to (4) we use the least ineffective prenexation of (3), namely

$$\forall x \exists v \forall w \exists y (A_D(x; y) \rightarrow B_D(v; w)) \quad (5)$$

and then apply the axiom of choice to obtain (4). This prenexation is only valid in the presence of semi-intuitionistic principles which turn out to be sound for the Dialectica interpretation. But again we have an asymmetry between premise and conclusion, as the functional  $f$  has access to both arguments  $x, w$  while the forward functional  $g$  only accesses  $x$ . As we will see, this is again a consequence of contraction being allowed in the premise but not in the conclusion (cf. also Shirahata's [30] discussion).

Finally there is the issue that the Dialectica interpretation requires decidability of quantifier-free formulas. This is once more due to the fact that in the rule of contraction, which can be written as  $A \rightarrow A \wedge A$ , we must produce a single negative move  $y$  in the game  $A_D(x; y)$  given two candidate negative moves  $y_0$  and  $y_1$ . By requiring that the adjudication relation  $A_D(x; y)$  is *decidable* we can simply check which of the two candidates  $y_0$  and  $y_1$  is actually the best negative move. The assumption of decidability, however, can be quite strong when working with higher order objects, and leads to restrictions on the amount of extensionality one is allowed to use.

As we have indicated, the root cause of all the asymmetry, obscurity and subtlety of the Dialectica interpretation comes from the need to deal with the naive-looking (and semantically trivial) structural rule of *contraction* (2). As we will discuss in the paper, this becomes even more evident once one moves to linear logic where contraction can be isolated from the other connectives. Without contraction, a formula  $A$  is actually interpreted in a much more symmetric way as

$$\left( \begin{array}{c} \exists x \\ \forall y \end{array} \right) A_D(x; y), \quad (6)$$

with the help of a simple form branching quantifier. For simplicity of notation we write this branching quantifier as  $\exists_y^x A$ . Without contraction, the intuitionistic implication  $A \rightarrow B$  become *linear implications*  $A \multimap B$ , and that can also be given a symmetric interpretation, since

$$\exists_y^x A_D(x; y) \multimap \exists_w^v B_D(v; w) \quad (7)$$

can be interpreted as

$$\exists_{x,w}^{f,g} (A_D(x; fw) \rightarrow B_D(gx; w)). \quad (8)$$

Finally, without contraction one does not need to assume decidability of quantifier-free formulas!

Unfortunately, one does need contraction in practise, since proofs often make use of a single assumption several times. In linear logic contraction is recovered in a controlled manner with the help of the modality  $!A$ . Contraction is only permitted for “marked” formulas  $!A$  and the rule (2) becomes

$$\frac{\Gamma, !A, !A \vdash B}{\Gamma, !A \vdash B} \text{ (con)}$$

We then have to give an interpretation to the new modality  $!A$ . As one might suspect, the interpretation of the modality introduces precisely the breaking of symmetry between the players. The new modality turns a symmetric game between Eloise and Abelard into a game where Abelard has the advantage of playing second and choosing his move based on Eloise’s move. One way of giving Abelard this advantage is to say that for the game  $!A$  Abelard’s move is a functional  $f$  which produces his move in game  $A$  given Eloise’s move. The adjudication relation for  $!A$  is then

$$(!A)_D(\mathbf{x}; f) \equiv A_D(\mathbf{x}; f\mathbf{x}).$$

For instance, the game associated with the formula  $A \equiv \exists_y^x(x \geq y)$  corresponds to the “biggest number game” in which both players try to cook up a bigger number than their opponent (equal numbers favour Eloise). Neither player has a winning move in this game. On the other hand, in the game  $!A$  Abelard has a winning move, since Eloise will have to provide an  $x$  and Abelard is asked to produce a function  $f$  such that  $x \geq fx$  is false, and he can choose e.g.  $f(x) = x + 1$ .

In linear logic the intuitionistic implication is derived via the linear implication  $A \multimap B$  and the modality  $!A$ , which does the bookkeeping of contractions, as

$$A \rightarrow B \equiv !A \multimap B. \quad (9)$$

Using this analysis of the intuitionistic implication we will see that the intrinsic difficulty of the Dialectica interpretation of  $A \rightarrow B$  discussed above comes from the subtle interpretation of  $!A$ .

The article is organised as follows. The next two subsections 1.1 and 1.2 give a brief introduction to Girard’s linear logic [11, 12] and Gödel’s Dialectica interpretation [1, 14]. Section 2 presents de Paiva’s Dialectica interpretation of linear logic. The relation between the Dialectica interpretations of linear logic and intuitionistic logic is discussed in Section 3. In Section 4 the characterisation principles required for the Dialectica interpretation of intuitionistic logic are analysed in the linear logic context. In Sections 5.1, 5.2 and 5.3 we look at extensions of Dialectica interpretation to classical logic, arithmetic and analysis, respectively. Finally, in Section 6 we discuss the relation between three variants of the Dialectica interpretation.

### 1.1 Intuitionistic linear logic

In the following we will describe a fragment of intuitionistic linear logic which is sufficient for embedding full intuitionistic logic. We work with an extension of the language

$\frac{\Gamma \vdash A \quad \Delta, A \vdash B}{\Gamma, \Delta \vdash B}$ (cut)	$\frac{\Gamma \vdash A}{\pi\{\Gamma\} \vdash A}$ (per)	$A_{\text{at}} \vdash A_{\text{at}}$ (id)
$\frac{\Gamma, A \vdash B}{\Gamma \vdash A \multimap B}$ ( $\multimap_r$ )	$\frac{\Gamma \vdash A}{\Gamma \vdash \forall z A}$ ( $\forall_r$ )	$\forall z A \vdash A[t/z]$ ( $\forall_l$ )
$\frac{\Gamma \vdash A \quad \Delta, B \vdash C}{\Gamma, \Delta, A \multimap B \vdash C}$ ( $\multimap_l$ )	$\frac{\Gamma, A \vdash B}{\Gamma, \exists z A \vdash B}$ ( $\exists_l$ )	$A[t/z] \vdash \exists z A$ ( $\exists_r$ )
$\frac{\Gamma[\gamma_0] \vdash A \quad \Gamma[\gamma_1] \vdash B}{\Gamma[(z)(\gamma_0, \gamma_1)] \vdash A \diamond_z B}$ ( $\diamond_z^l$ )		$A \vdash A \diamond_t B$ $A \diamond_t B \vdash A$ } ( $\diamond_t$ )
$\frac{\Gamma[\gamma_0], A \vdash C[\mathbf{c}_0] \quad \Gamma[\gamma_1], B \vdash C[\mathbf{c}_1]}{\Gamma[(z)(\gamma_0, \gamma_1)], A \diamond_z B \vdash C[(z)(\mathbf{c}_0, \mathbf{c}_1)]}$ ( $\diamond_z^r$ )		$B \vdash A \diamond_f B$ $A \diamond_f B \vdash B$ } ( $\diamond_f$ )

Table 2. Exponential-free intuitionistic linear logic

of linear logic to all finite types. The set of *finite types*  $\mathbb{T}$  is inductively defined as follows:  $b, i \in \mathbb{T}$ ; and if  $\rho, \sigma \in \mathbb{T}$  then  $\rho \rightarrow \sigma \in \mathbb{T}$ . For simplicity, we work with just two basic finite types  $b$  (boolean) and  $i$  (integer).

The terms of the language contain all typed  $\lambda$ -terms, i.e. variables  $x^\rho$  for each finite type  $\rho$ ;  $\lambda$ -abstractions  $(\lambda x^\rho. t^\sigma)^{\rho \rightarrow \sigma}$ ; and term applications  $(t^{\rho \rightarrow \sigma} s^\rho)^\sigma$ . Besides these we also add boolean constants for true and false, and the if-then-else term construction  $(t^b)(r^\tau, s^\tau)$  for each type  $\tau$ . The term  $(t^b)(r^\tau, s^\tau)$  reduces to either  $r$  or  $s$  depending on whether the boolean term  $t$  reduces to true or false, respectively.

The atomic formulas are  $A_{\text{at}}, B_{\text{at}}, \dots$ . For simplicity, the standard propositional constants of intuitionistic linear logic have been omitted, since the interpretation of atomic formulas is trivial (see Section 2). Formulas are built out of atomic formulas  $A_{\text{at}}, B_{\text{at}}, \dots$  via the connectives  $A \multimap B$  (linear implication),  $A \diamond_z B$  (if-then-else), and quantifiers  $\forall x A$  and  $\exists x A$ . The rules for these are shown in Table 2, with the usual side condition in the rules ( $\forall_r$ ) and ( $\exists_l$ ) that the variable  $z$  must not appear free in  $\Gamma, B$ . The structural rules of linear logic do not contain the usual rules of weakening and contraction. These are added separately in a controlled manner via the use of the modality  $!A$ . The rules governing the behaviour of  $!A$  are shown in Table 3.

Note that we are deviating from the standard formulation of linear logic, in the sense that we will use the if-then-else logical constructor  $A \diamond_z B$  instead of standard additive conjunction and disjunction<sup>1</sup>. The standard additive connectives can be defined as

<sup>1</sup> See Girard's comments in [11] (p13) and [12] (p73) on the relation between the additive connectives and the if-then-else construct.

$\frac{! \Gamma \vdash A}{! \Gamma \vdash ! A} (!_r)$	$\frac{\Gamma, A \vdash B}{\Gamma, ! A \vdash B} (!_l)$	$\frac{\Gamma, ! A, ! A \vdash B}{\Gamma, ! A \vdash B} (\text{con})$	$\frac{\Gamma \vdash B}{\Gamma, ! A \vdash B} (\text{wkn})$
---	---	---	---

Table 3. Rules for the exponentials

$$A \wedge B := \forall z^b (A \diamond_z B)$$

$$A \vee B := \exists z^b (A \diamond_z B)$$

with the help of quantification over booleans. In the following we sometimes use the abbreviations  $A \wedge B$  and  $A \vee B$  to stand for the corresponding formulas containing the if-then-else connective.

Since we will be considering extensions of this basic system, let us denote the system just described by  $\text{LL}_0^\omega$ . After introducing the basic Dialectica interpretation of  $\text{LL}_0^\omega$  we will discuss extensions of this basic system for which the interpretation can be characterised.

**Notation.** We use bold face variables  $\mathbf{f}, \mathbf{g}, \dots, \mathbf{x}, \mathbf{y}, \dots$  for tuples of variables, and bold face terms  $\mathbf{a}, \mathbf{b}, \dots, \gamma, \delta, \dots$  for tuples of terms. Given sequences of terms  $\mathbf{a}$  and  $\mathbf{b}$ , by  $\mathbf{a}(\mathbf{b})$  we mean the sequence of terms  $a_0(\mathbf{b}), \dots, a_n(\mathbf{b})$ ; and by  $\mathbf{a}[\mathbf{b}/\mathbf{x}]$  we mean the sequence  $a_0[\mathbf{b}/\mathbf{x}], \dots, a_n[\mathbf{b}/\mathbf{x}]$ .

## 1.2 Gödel's Dialectica interpretation

Gödel's Dialectica interpretation [1, 7, 14] is normally presented as in Definition 1 below. Note the asymmetric treatment of conjunction/disjunction and universal/existential quantifiers.

**Definition 1 (Dialectica interpretation).** For each formula  $A$  of intuitionistic logic we associate a new quantifier-free formula  $A_D(\mathbf{x}; \mathbf{y})$  inductively as follows:

$$(A_{\text{at}})^D := A_{\text{at}}, \quad \text{when } A_{\text{at}} \text{ is an atomic formula.}$$

Assume we have already defined  $A_D(\mathbf{x}; \mathbf{y})$  and  $B_D(\mathbf{v}; \mathbf{w})$ . We then define

$$(A \wedge B)_D(\mathbf{x}, \mathbf{v}; \mathbf{y}, \mathbf{w}) := A_D(\mathbf{x}; \mathbf{y}) \wedge B_D(\mathbf{v}; \mathbf{w})$$

$$(A \vee B)_D(\mathbf{x}, \mathbf{v}, z; \mathbf{y}, \mathbf{w}) := A_D(\mathbf{x}; \mathbf{y}) \diamond_z B_D(\mathbf{v}; \mathbf{w})$$

$$(A \rightarrow B)_D(\mathbf{f}, \mathbf{g}; \mathbf{x}, \mathbf{w}) := A_D(\mathbf{x}; \mathbf{f}\mathbf{w}\mathbf{x}) \rightarrow B_D(\mathbf{g}\mathbf{x}; \mathbf{w})$$

$$(\forall z A)_D(\mathbf{f}; \mathbf{y}, z) := A_D(\mathbf{f}z; \mathbf{y})$$

$$(\exists z A)_D(\mathbf{x}, z; \mathbf{y}) := A_D(\mathbf{x}; \mathbf{y}).$$

Finally, we define  $(A)^D := \exists \mathbf{x} \forall \mathbf{y} A_D(\mathbf{x}; \mathbf{y})$ .

In the following, we will work with an equivalent formulation of Gödel's Dialectica interpretation where the clauses for conjunction, disjunction and the existential quantifiers are slightly modified, so as to treat these in a symmetric way. For the sake of reference we state this variant as a definition:

**Definition 2 (Equivalent formulation of Dialectica interpretation).** *Same as in Definition 1 except that the treatment of disjunction, conjunction and the existential quantifier are modified as:*

$$\begin{aligned} (A \wedge B)_D(\mathbf{x}, \mathbf{v}; \mathbf{y}, \mathbf{w}, z^b) &::= A_D(\mathbf{x}z; \mathbf{y}) \diamond_z B_D(\mathbf{v}z; \mathbf{w}) \\ (A \vee B)_D(\mathbf{x}, \mathbf{v}, z^b; \mathbf{y}, \mathbf{w}) &::= A_D(\mathbf{x}; \mathbf{y}z) \diamond_z B_D(\mathbf{v}; \mathbf{w}z) \\ (\exists z^\tau A)_D(\mathbf{x}, z; \mathbf{f}) &::= A_D(\mathbf{x}; \mathbf{f}z). \end{aligned}$$

The formula  $(A)^D$  is defined as before, i.e.  $(A)^D ::= \exists \mathbf{x} \forall \mathbf{y} A_D(\mathbf{x}; \mathbf{y})$ .

This reformulation leads to (intuitionistically) equivalent formulas  $(A)^D$ . In the case of conjunction and disjunction, the extra boolean information  $z$  given to the functionals is irrelevant, since each functional will only be applied when the boolean is either true or false. The reason for allowing the boolean  $z$  as an argument for the functionals will become clear in Section 3, where we show how this comes naturally from the interpretation of linear logic. In general terms, we can argue for the equivalence between the two choices of  $A^D$  by noticing that both can be shown to be equivalent to  $A$  using the same characterisation principles (see Section 4). Since equivalences between  $\Sigma_2$ -statements shown using the characterisation principles can also be shown without these principles the result then follows. The benefit of modifying the interpretation is that we obtain a symmetric treatment of the connectives  $(\vee/\wedge)$  and quantifiers  $(\exists/\forall)$ . This symmetry, which in this case is optional, is forced upon us in the case of the Dialectica interpretation of linear logic, as we will see in the following section.

## 2 The Dialectica Interpretation of Linear Logic

In this section we present de Paiva's Dialectica interpretation [25] of intuitionistic linear logic. To each formula  $A$  of the fragment of linear logic  $\text{LL}_0^\omega$  we associate a quantifier-free formula  $|A|_{\mathbf{y}}^{\mathbf{x}}$ , where  $\mathbf{x}, \mathbf{y}$  are fresh-variables not appearing in  $A$ . The variables  $\mathbf{x}$  in the superscript are called the *witnessing variables*, while the subscript variables  $\mathbf{y}$  are called the *challenge variables*. Intuitively, the interpretation of a formula  $A$  is a one-move two-player (Eloise and Abelard) game, where  $|A|_{\mathbf{y}}^{\mathbf{x}}$  is the adjudication relation. We want Eloise to have a winning move whenever  $A$  is provable in  $\text{LL}_0^\omega$ . Consider first the case of linear implication  $A \multimap B$ . In this game, Eloise claims to have a justification for  $B$  given a justification for  $A$ , and also, claims to give a refutation for  $A$  given a refutation for  $B$ . Hence, her move in this game is a pair of constructions  $\mathbf{f}, \mathbf{g}$ , while Abelard must present arguments  $\mathbf{x}$  in favour of  $A$  and  $\mathbf{w}$  against  $B$ . Using our notation the adjudication relation for the game  $A \multimap B$  can be succinctly described as

$$|A \multimap B|_{\mathbf{x}, \mathbf{w}}^{\mathbf{f}, \mathbf{g}} ::= |A|_{\mathbf{f}\mathbf{w}}^{\mathbf{x}} \multimap |B|_{\mathbf{w}}^{\mathbf{g}\mathbf{x}}.$$

The if-then-else game is just a flagged disjoint union of the two games  $A$  and  $B$ , where the boolean flag  $z$  tells which game is indeed being played. Therefore, both players make moves in both games  $A$  and  $B$ , but only one of the games will be considered depending on the boolean value of  $z$ , i.e.

$$|A \diamond_z B|_{\mathbf{y}, \mathbf{w}}^{\mathbf{x}, v} := |A|_{\mathbf{y}}^{\mathbf{x}} \diamond_z |B|_{\mathbf{w}}^v.$$

The quantifier games correspond to a parametrised family of games. Depending on whether we have a universal or existential quantifier, one of the players selects which particular instance  $z$  of the game they want to play, while the other player is allowed to choose their move depending on the choice  $z$  of the instantiation. For instance, in the case of the game  $\exists z A$ , Eloise chooses a value for  $z$  and makes a move  $\mathbf{x}$  in the game  $A(z)$ . Abelard's move is a functional  $f$  which transforms Eloise's choice of instantiation  $z$  into his move  $fz$ . A symmetric situation occurs in the game  $\forall z A$ , i.e.

$$|\exists z A(z)|_{\mathbf{f}}^{\mathbf{x}, z} := |A(z)|_{\mathbf{f}z}^{\mathbf{x}}$$

$$|\forall z A(z)|_{\mathbf{y}, z}^{\mathbf{f}} := |A(z)|_{\mathbf{y}}^{\mathbf{f}z}.$$

In the game  $!A$  the symmetry between the players is broken. Eloise must make her move  $\mathbf{x}$  available to Abelard, and depending on  $\mathbf{x}$  Abelard chooses his move. Equivalently, we ask that Abelard's move in this game be a functional  $f$  producing his move whenever given Eloise's move, i.e.

$$|!A|_{\mathbf{f}}^{\mathbf{x}} := |!A|_{\mathbf{f}\mathbf{x}}^{\mathbf{x}}.$$

The interpretation of linear logic formulas given above is such that a linear logic proof of  $A$  will provide Eloise's winning move  $\mathbf{a}$  in the corresponding game, i.e.  $\forall \mathbf{y} |A|_{\mathbf{y}}^{\mathbf{a}}$ . This is formally stated in the following theorem.

**Theorem 1 (Soundness).** *Let  $A_0, \dots, A_n, B$  be formulas of  $\text{LL}_0^\omega$ , with  $z$  as the only free-variables. If*

$$A_0(z), \dots, A_n(z) \vdash B(z)$$

*is provable in  $\text{LL}_0^\omega$  then terms  $\mathbf{a}_0, \dots, \mathbf{a}_n, \mathbf{b}$  can be extracted from this proof such that*

$$|A_0(z)|_{\mathbf{a}_0}^{\mathbf{x}_0}, \dots, |A_n(z)|_{\mathbf{a}_n}^{\mathbf{x}_n} \vdash |B(z)|_{\mathbf{w}}^{\mathbf{b}}$$

*is provable in the quantifier-free fragment of  $\text{LL}_0^\omega$ , where*

- $\text{FV}(\mathbf{a}_i) \in \{z, \mathbf{w}, \mathbf{x}_0, \dots, \mathbf{x}_n\} \setminus \{\mathbf{x}_i\}$
- $\text{FV}(\mathbf{b}) \in \{z, \mathbf{x}_0, \dots, \mathbf{x}_n\}$ .

**Proof.** See [22–24] for details of how the terms  $\mathbf{a}_0, \dots, \mathbf{a}_n, \mathbf{b}$  and a derivation of

$$|A_0(z)|_{\mathbf{a}_0}^{\mathbf{x}_0}, \dots, |A_n(z)|_{\mathbf{a}_n}^{\mathbf{x}_n} \vdash |B(z)|_{\mathbf{w}}^{\mathbf{b}}$$

can be constructed by induction on the derivation of  $A_0(z), \dots, A_n(z) \vdash B(z)$ .  $\square$

*Remark 1 (Chu spaces).* One can also view the Dialectica interpretation above as associating formulas  $A$  of  $\text{LL}_0^\omega$  with Chu spaces, i.e. triples  $(\rho, \tau, |A|_{\mathbf{y}}^{\mathbf{x}})$ , where  $|A|_{\mathbf{y}}^{\mathbf{x}}$  is a relation between  $\mathbf{x} \in \rho$  and  $\mathbf{y} \in \tau$ . The Dialectica constructions, however, differ slightly from the Chu constructions, as discussed in de Paiva's recent paper [27].



The Dialectica interpretation of intuitionistic logic (described in Section 1.2) gives a strange advantage to Abelard, since he is only asked to present his move after Eloise has chosen  $\mathbf{x}$ . In view of the Dialectica interpretation of linear logic presented above, this lack of symmetry can be seen to come from the fact that formulas of intuitionistic logic are viewed as  $!A$ , and the modality  $!$  indeed breaks the symmetry between the two players. If one looks at the interpretation of linear logic presented above, however, except for the treatment of  $!A$ , there is a nice conformity between the two players. The reason is that a formula  $A$  of linear logic is actually interpreted as

$$\exists_y^x |A|_y^x \quad (10)$$

where  $\exists_y^x A$  is a simple form of branching quantifier (we will refer to these as *simultaneous quantifiers*) for which we assume the following rule:

$$\frac{A_0(\mathbf{x}_0, \mathbf{a}_0), \dots, A_n(\mathbf{x}_n, \mathbf{a}_n) \vdash B(\mathbf{b}, \mathbf{w})}{\exists_{\mathbf{y}_0}^{\mathbf{x}_0} A_0(\mathbf{x}_0, \mathbf{y}_0), \dots, \exists_{\mathbf{y}_n}^{\mathbf{x}_n} A_n(\mathbf{x}_n, \mathbf{y}_n) \vdash \exists_{\mathbf{w}}^v B(\mathbf{v}, \mathbf{w})} (\exists')$$

with the two side-conditions:

- $\mathbf{x}_i$  may only appear free in the terms  $\mathbf{b}$  or  $\mathbf{a}_j$ , for  $j \neq i$ ;
- $\mathbf{w}$  may only appear free in the terms  $\mathbf{a}_i$ .

The simultaneous quantifier subsumes both existential and universal quantifiers, when the tuples  $\mathbf{w}$  and  $\mathbf{u}$  are empty in  $\exists_{\mathbf{w}}^v B(\mathbf{v}, \mathbf{w})$ , respectively. In terms of games, the new quantifier embodies the idea of the two players performing their moves simultaneously. Let us refer to the extension of  $\text{LL}_0^\omega$  with the new simultaneous quantifier by  $\text{LL}_1^\omega$ .

**Theorem 2.** *Extend the interpretation above to the system  $\text{LL}_1^\omega$  by defining*

$$|\exists_{\mathbf{w}}^v A(\mathbf{v}, \mathbf{w})|_{\mathbf{g}, \mathbf{w}}^{\mathbf{f}, \mathbf{v}} := |A(\mathbf{v}, \mathbf{w})|_{\mathbf{g}, \mathbf{v}}^{\mathbf{f}, \mathbf{w}}.$$

*Theorem 1 holds for the extended system  $\text{LL}_1^\omega$ .*

**Proof.** We must show that the soundness theorem still holds when the system is extended with the new rule for the simultaneous quantifier. The new rule is handled as follows:

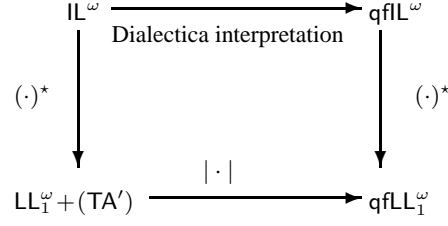
$$\frac{\frac{|A_0(\mathbf{v}_0, \mathbf{a}_0)|_{\gamma_0}^{\mathbf{x}_0}, \dots, |A_n(\mathbf{v}_0, \mathbf{a}_n)|_{\gamma_n}^{\mathbf{x}_n} \vdash |B(\mathbf{b}, \mathbf{y})|_{\mathbf{u}}^{\delta}}{|A_0(\mathbf{v}_0, \mathbf{a}_0)|_{\gamma_0'}^{\mathbf{f}_0, \mathbf{a}_0}, \dots, |A_n(\mathbf{v}_0, \mathbf{a}_n)|_{\gamma_n'}^{\mathbf{f}_n, \mathbf{a}_n} \vdash |B(\mathbf{b}, \mathbf{y})|_{\mathbf{g}, \mathbf{b}}^{\delta'} \left[ \frac{\mathbf{f}_i, \mathbf{a}_i}{\mathbf{x}_i}, \frac{\mathbf{g}, \mathbf{b}}{\mathbf{u}} \right]}}{|\exists_{\mathbf{w}_0}^{\mathbf{v}_0} A_0(\mathbf{v}_0, \mathbf{w}_0)|_{\gamma_0'}^{\mathbf{f}_0, \mathbf{v}_0}, \dots, |\exists_{\mathbf{w}_n}^{\mathbf{v}_n} A_n(\mathbf{v}_n, \mathbf{w}_n)|_{\gamma_n'}^{\mathbf{f}_n, \mathbf{v}_n} \vdash |\exists_{\mathbf{y}}^x B(\mathbf{x}, \mathbf{y})|_{\mathbf{g}, \mathbf{y}}^{\delta', \mathbf{b}}}$$

where

- $\delta' := \lambda \mathbf{y}. \delta[\mathbf{f}_i, \mathbf{a}_i / \mathbf{x}_i]$ , for all  $i$
- $\gamma_j' := \lambda \mathbf{v}_j. \gamma_j[\mathbf{g}, \mathbf{b} / \mathbf{u}][\mathbf{f}_i, \mathbf{a}_i / \mathbf{x}_i]$ , for all  $i \neq j$ .

This concludes the proof.  $\square$

In fact, note that the simultaneous quantifiers are eliminated via the interpretation, and we obtain an interpretation of  $\text{LL}_1^\omega$  into the quantifier-free fragment of  $\text{LL}_0^\omega$ . Theorem 2 also provides an assurance that the rule for the simultaneous quantifier suggested above characterises that quantifier, since a proof of the conclusion of the rule yields terms satisfying the premise.



**Fig. 1.** Composing Girard's embedding with Dialectica interpretation of  $\text{LL}_1^\omega + (\text{TA}')$

### 3 Relation to Dialectica Interpretation of Intuitionistic Logic

In this section we describe how the interpretation of linear logic presented above corresponds to Gödel's Dialectica interpretation [14] of intuitionistic logic. A similar result concerning the Dialectica interpretation of classical logic is shown in Shirahata [30]. We will assume that intuitionistic logic is also formalised with the if-then-else connective  $A \diamond_z B$ , so that conjunction and disjunction are defined notions. First, consider a variant of Girard's embedding of intuitionistic logic into our version of linear logic with conditionals.

**Definition 3 ([11]).** For any formula  $A$  of intuitionistic logic its linear translation  $A^*$  is defined inductively as

$$\begin{aligned}
(A_{\text{at}})^* &::= A_{\text{at}} \\
(A \diamond_z B)^* &::= A^* \diamond_z B^* \\
(A \rightarrow B)^* &::= !A^* \multimap B^* \\
(\forall z A)^* &::= \forall z A^* \\
(\exists z A)^* &::= \exists z A^*.
\end{aligned}$$

Girard's original embedding makes use of an extra  $!$  for the interpretation of  $\exists x A$  as  $\exists x !A^*$ . We will see that in the presence of the principle

$$(\text{TA}') \quad !\exists z A \multimap \exists z !A.$$

this extra  $!$  is not necessary. The reason for assuming the principle  $(\text{TA}')$  is that it is sound for the Dialectica interpretation under consideration. In fact, this is a simple form of the more general principle  $(\text{TA})$  to be discussed in Section 4.4.

**Theorem 3.** Let  $\Gamma, A$  be formulas in the language of intuitionistic logic. The translation given in Definition 3 is such that if  $\Gamma \vdash A$  is derivable in  $\mathbb{L}^\omega$  then  $!(\Gamma^*) \vdash A^*$  is derivable in  $\text{LL}_1^\omega + (\text{TA}')$ .

**Proof.** By induction on the derivation of  $\Gamma \vdash A$  in  $\mathbb{L}^\omega$ . The only difference from Girard's embedding is in the treatment of the existential quantifier. With the help of the

extra principle (TA') it can be treated as

$$\frac{\frac{\frac{! \Gamma^*, ! A^* \vdash B^*}{! \Gamma^*, \exists z ! A^* \vdash B^*} (\exists_l)}{! \Gamma^*, ! \exists z A^* \vdash B^*} (\text{TA}')}{! \Gamma^*, ! (\exists z A)^* \vdash B^*} (\text{D3})$$

The other cases are treated as usual (see [28]).  $\square$

The next theorem states that, up to the embedding described in Def. 3, the Dialectica interpretation of intuitionistic logic corresponds to the interpretation of linear logic (see diagram of Figure 1).

**Theorem 4.** *Let  $A$  be a formula of intuitionistic logic. Then  $(A_D(\mathbf{x}; \mathbf{y}))^* \equiv |A^*|_{\mathbf{y}}^{\mathbf{x}}$  (where  $\equiv$  denotes syntactic identity).*

**Proof.** The proof is by induction on the logical structure of the intuitionistic formula  $A$ . Recall that  $A \vee B$  is defined as  $\exists z(A \diamond_z B)$ . We have

*Disjunction*

$$\begin{aligned} ((A \vee B)_D(\mathbf{x}, \mathbf{v}, z; \mathbf{y}, \mathbf{w}))^* &\stackrel{(\text{Sec1.2})}{\equiv} (A_D(\mathbf{x}; \mathbf{y}z) \diamond_z B_D(\mathbf{v}; \mathbf{w}z))^* \\ &\stackrel{(\text{Def3})}{\equiv} (A_D(\mathbf{x}; \mathbf{y}z))^* \diamond_z (B_D(\mathbf{v}; \mathbf{w}z))^* \\ &\stackrel{(\text{IH})}{\equiv} |A^*|_{\mathbf{y}z}^{\mathbf{x}} \diamond_z |B^*|_{\mathbf{w}z}^{\mathbf{v}} \stackrel{(\text{Sec2})}{\equiv} |\exists z(A \diamond_z B^*)|_{\mathbf{y}, \mathbf{w}}^{\mathbf{x}, \mathbf{v}, z} \\ &\stackrel{(\text{Def3})}{\equiv} |(A \vee B)^*|_{\mathbf{y}, \mathbf{w}}^{\mathbf{x}, \mathbf{v}, z}, \end{aligned}$$

*Implication*

$$\begin{aligned} ((A \rightarrow B)_D(\mathbf{f}, \mathbf{g}; \mathbf{x}, \mathbf{w}))^* &\stackrel{(\text{Sec1.2})}{\equiv} (A_D(\mathbf{x}; \mathbf{g}\mathbf{x}\mathbf{w}) \rightarrow B_D(\mathbf{f}\mathbf{x}; \mathbf{w}))^* \\ &\stackrel{(\text{Def3})}{\equiv} !(A_D(\mathbf{x}; \mathbf{g}\mathbf{x}\mathbf{w}))^* \multimap (B_D(\mathbf{f}\mathbf{x}; \mathbf{w}))^* \\ &\stackrel{(\text{IH})}{\equiv} !|A^*|_{\mathbf{g}\mathbf{x}\mathbf{w}}^{\mathbf{x}} \multimap |B^*|_{\mathbf{w}}^{\mathbf{f}\mathbf{x}} \\ &\stackrel{(\text{Sec2})}{\equiv} !|A^*|_{\mathbf{g}\mathbf{w}}^{\mathbf{x}} \multimap |B^*|_{\mathbf{w}}^{\mathbf{f}\mathbf{x}} \stackrel{(\text{Sec2})}{\equiv} !|A^* \multimap B^*|_{\mathbf{x}, \mathbf{w}}^{\mathbf{f}, \mathbf{g}} \\ &\stackrel{(\text{Def3})}{\equiv} |(A \rightarrow B)^*|_{\mathbf{x}, \mathbf{w}}^{\mathbf{f}, \mathbf{g}}. \end{aligned}$$

*Existential quantifier*

$$\begin{aligned} ((\exists z A)_D(\mathbf{x}, z; \mathbf{f}))^* &\stackrel{(\text{Sec1.2})}{\equiv} (A_D(\mathbf{x}; \mathbf{f}z))^* \stackrel{(\text{IH})}{\equiv} |A^*|_{\mathbf{f}z}^{\mathbf{x}} \stackrel{(\text{Sec2})}{\equiv} |\exists z A^*|_{\mathbf{f}}^{\mathbf{x}, z} \\ &\stackrel{(\text{Def3})}{\equiv} |(\exists z A)^*|_{\mathbf{f}}^{\mathbf{x}, z}, \end{aligned}$$

The treatment of conjunction and universals is similar to disjunction and existentials, respectively.  $\square$

In the remaining sections of the paper we will consider further extensions of the Dialectica interpretation of linear logic, and show how these extensions correspond to extensions of Gödel's Dialectica interpretation. We start by considering the principles needed for the characterisation of the interpretation, then we conclude with the extensions covering classical logic, arithmetic and analysis.

## 4 An Analysis of the Characterisation Principles

As mentioned in Section 1.2, the Dialectica interpretation is based on a prenexation of an arbitrary formula  $A$  into a formula of the form (1). The interpretation is such that an intuitionistic proof of  $A$  gives us enough information to explicitly construct a witness for the existential quantifier of (1). In fact, the following semi-intuitionistic principles: *axiom of choice for universal formulas*

$$(AC) \quad \forall \mathbf{x} \exists \mathbf{y} A_{\forall}(\mathbf{x}, \mathbf{y}) \rightarrow \exists \mathbf{f} \forall \mathbf{x} A_{\forall}(\mathbf{x}, \mathbf{f} \mathbf{x})$$

*independence of universal premises*

$$(IP) \quad (A_{\forall} \rightarrow \exists \mathbf{x} B_{\forall}) \rightarrow \exists \mathbf{x} (A_{\forall} \rightarrow B_{\forall})$$

and *Markov principle*

$$(MP) \quad (\forall \mathbf{x} A_{\text{qf}} \rightarrow B_{\text{qf}}) \rightarrow \exists \mathbf{x} (A_{\text{qf}} \rightarrow B_{\text{qf}})$$

are sufficient for showing the equivalence between  $A$  and its interpretation (1). Above, by  $A_{\text{qf}}, B_{\text{qf}}$  we mean quantifier-free formulas, whereas  $A_{\forall}, B_{\forall}$  denote purely universal formulas, i.e. formulas of the form  $\forall z A_{\text{qf}}(z)$ .

Let us look at each of the principles in turn and investigate their linear logic counterparts. In other words, we look at which linear logic principles to use in order to derive the linear logic translations  $(AC)^*$ ,  $(IP)^*$  and  $(MP)^*$ . These will allow us to derive the equivalence between an arbitrary formula  $A$  and its (linear) Dialectica interpretation  $\exists \mathbf{y}^x |A|_{\mathbf{y}}^x$ .

### 4.1 Axiom of choice

The axiom of choice (AC) says that a  $\forall \exists$  quantifier dependence implies the existence of a functional realising this dependence. When we move into the linear logic context, the axiom of choice (AC) boils down to the fact that if Eloise wins a parametrised game  $\exists \mathbf{y}^x A_{\text{qf}}(\mathbf{x}, \mathbf{y}, z)$  for any choice of  $z$ , then she must have a strategy  $\mathbf{f}$  which produces her moves whenever given the value of the parameter  $z$ , i.e.

$$(AC_l) \quad \forall z \exists \mathbf{y}^x A_{\text{qf}}(\mathbf{x}, \mathbf{y}, z) \multimap \exists \mathbf{f}_{\mathbf{y}, z}^x A_{\text{qf}}(\mathbf{f} z, \mathbf{y}, z).$$

Moreover, since we are working with linear implication, its contrapositive

$$(AC_l^c) \quad \exists \mathbf{f}^{x, z} A_{\text{qf}}(\mathbf{x}, \mathbf{f} z, z) \multimap \exists z \exists \mathbf{y}^x A_{\text{qf}}(\mathbf{x}, \mathbf{y}, z)$$

is also realised by the linear Dialectica interpretation. Note that the converse of both  $(AC_l)$  and  $(AC_l^c)$  can be derived in  $\text{LL}_1^\omega$ . Assuming that  $A$  is equivalent to  $\exists \mathbf{y}^x |A|_{\mathbf{y}}^x$ , one then obtains the equivalence between  $\forall z A$  or  $\exists z A$  and their interpretations as:

$$\begin{aligned} \forall z A &\stackrel{(\text{IH})}{\Leftrightarrow} \forall z \exists \mathfrak{Y}_y^x |A|_y^x \stackrel{(\text{AC}_I, \text{LL}_1^\omega)}{\Leftrightarrow} \exists \mathfrak{Y}_{y,z}^f |A|_{fz}^f, \\ \exists z A &\stackrel{(\text{IH})}{\Leftrightarrow} \exists z \exists \mathfrak{Y}_y^x |A|_y^x \stackrel{(\text{LL}_1^\omega, \text{AC}_I^c)}{\Leftrightarrow} \exists \mathfrak{Y}_f^{y,z} |A|_{fz}^x. \end{aligned}$$

We will also make use of a simple form of  $(\text{AC}_I^c)$  in Section 4.4, for the treatment of the modality  $!A$ .

## 4.2 Independence of premise

The independence of premise says that universal assumptions do not contribute to witnessing existential statements. Existential quantifiers can then be prenexed over such universal premises. The linear logic counterpart of (IP) states that if Eloise wins a games  $\exists \mathfrak{Y}_w^v B_{\text{qf}}(v)$  relative to another game  $\exists \mathfrak{Y}_y^x A_{\text{qf}}(y)$  it must be because she has copy-cat strategies  $f, g$  which take positive moves in the first game into positive moves of the second, and negative moves of the second into negative moves of the first, i.e.

$$(\text{IP}_I) \quad (\exists \mathfrak{Y}_y^x A_{\text{qf}}(y) \multimap \exists \mathfrak{Y}_w^v B_{\text{qf}}(v)) \multimap \exists \mathfrak{Y}_{x,w}^{f,g} (A_{\text{qf}}(gw) \multimap B_{\text{qf}}(fx)).$$

This is in fact a generalisation of the independence of premise principle (case when tuples  $x$  and  $w$  are empty). Again, it is easy to see that the converse of  $(\text{IP}_I)$  is derivable in  $\text{LL}_1^\omega$ . Assuming that  $A$  and  $B$  are equivalent to  $\exists \mathfrak{Y}_y^x |A|_y^x$  and  $\exists \mathfrak{Y}_w^v |B|_w^v$ , respectively, one then obtains the equivalence between  $A \multimap B$  and its interpretation as:

$$A \multimap B \stackrel{(\text{IH})}{\Leftrightarrow} \exists \mathfrak{Y}_y^x |A|_y^x \multimap \exists \mathfrak{Y}_w^v |B|_w^v \stackrel{(\text{IP}_I, \text{LL}_1^\omega)}{\Leftrightarrow} \exists \mathfrak{Y}_{x,w}^{f,g} (|A|_{gw}^x \multimap |B|_{fw}^v).$$

## 4.3 Markov principle

As the independence of premise, the Markov principle (MP) is a classically valid principle which turns out to be validated by the Dialectica interpretation. Its importance comes from the fact that up to the Markov principle the negative translation of  $\Pi_2^0$  statements is intuitionistically equivalent to the statements themselves. This is crucial for concrete applications of proof theory to mathematics [17, 18, 29].

In the case of *intuitionistic logic*, the Markov principle is used to obtain the equivalence between  $\forall x A_{\text{qf}} \rightarrow B_{\text{qf}}$  and  $\exists x (A_{\text{qf}} \rightarrow B_{\text{qf}})$ . On the other hand, in the case of *linear logic*, given  $(\text{IP}_I)$ , all we need to add is the principle

$$(\text{MP}_I) \quad \forall x !A_{\text{qf}} \multimap !\forall x A_{\text{qf}}$$

establishing the commutativity between the universal quantifier and the ‘‘bang’’ modality. This equivalence is not derivable in standard linear logic, but is validated by the Dialectica interpretation. Having many copies of  $\forall x A_{\text{qf}}$  seems to be stronger than only being able to instantiate  $\forall x !A_{\text{qf}}$  once and then being able to use that instantiation several times. What the Dialectica interpretation shows is that this is not the case, at least when we have decidability of  $A_{\text{qf}}$ . The reason is that having decidability of quantifier-free formulas allows us to incorporate several instantiations into a single instantiation, via the definition by cases.

Using  $(\text{IP}_I)$  and  $(\text{MP}_I)$  we can derive  $(\text{MP})^*$  as

$$(!\forall x A_{\text{qf}} \multimap B_{\text{qf}}) \stackrel{(\text{MP}_I)}{\multimap} (\forall x !A_{\text{qf}} \multimap B_{\text{qf}}) \stackrel{(\text{IP}_I)}{\multimap} \exists x (!A_{\text{qf}} \multimap B_{\text{qf}}).$$

As it turns out, in order to show the equivalence between  $!A$  and its interpretation we need yet an extra principle, which is not visible in the intuitionistic context.

#### 4.4 An extra principle!

By refining the Dialectica interpretation via linear logic we also get soundness with respect to a new principles which, like  $(MP_l)$ , can not even be stated in the context of intuitionistic logic, namely

(TA)  $! \exists y^x A \multimap \exists x ! \forall y A$ .

We refer to this principle as *trump advantage*, since the “bang” modality turns a symmetric game  $\exists y^x A$  into a game where Eloise has to play first, and Abelard’s move is allowed to depend on Eloise’s. A simple form of this principle has already been discussed in Section 3, and basically states that the  $!$ -modality commutes with the existential quantifier.

Assuming that  $A$  is equivalent to  $\exists y^x |A|_y^x$ , we are then able to derive the equivalence between  $!A$  and its interpretation as:

$$\begin{array}{ccccccc}
 !A & \xleftrightarrow{\text{IH}} & !\exists y^x |A|_y^x & \xrightarrow{\text{TA}} & \exists x ! \forall y |A|_y^x & \xrightarrow{\text{LL}_0^\omega} & \exists x \forall y ! |A|_y^x & \xrightarrow{\text{LL}_1^\omega} & \exists y^x ! |A|_y^x \\
 & & & & \xleftarrow{\text{LL}_1^\omega} & & \xleftarrow{\text{MP}_l} & & \xleftarrow{\text{AC}_l^c}
 \end{array}$$

#### 4.5 Characterisation of linear logic Dialectica

Let us denote by  $\text{LL}_2^\omega$  the extension of  $\text{LL}_1^\omega$  with these four schemata  $(AC_l)$ ,  $(IP_l)$ ,  $(MP_l)$  and  $(TA)$ . The next lemma states that, in fact, these extra principles are all one needs to show the equivalence between  $A$  and its interpretation  $\exists y^x |A|_y^x$ .

**Lemma 1.** *The equivalence between  $A$  and  $\exists y^x |A|_y^x$  can be derived in the system  $\text{LL}_2^\omega$ .*

**Proof.** This has been shown above.  $\square$

In particular, we obtain a (weak) completeness result, showing how winning moves for Eloise in a game  $|A|_y^x$  correspond to the linear logic formula  $A$  being provable.

**Theorem 5.** *Let  $A$  be a formula in the language of  $\text{LL}_0^\omega$ . Then  $A$  is derivable in  $\text{LL}_2^\omega$  if and only if  $|A|_y^t$  is derivable in  $\text{LL}_0^\omega$ , for some sequence of terms  $t$ .*

**Proof.** The forward implication follows from the extensions of the Soundness Theorem 1 discussed above. For the other direction, assume  $|A|_y^t$  is derivable in  $\text{LL}_0^\omega$ . In particular we have  $\exists y^x |A|_y^x$ . By Lemma 1 we get that  $A$  is derivable in  $\text{LL}_2^\omega$ .  $\square$

## 5 Extensions of the Dialectica Interpretation

In this section we analyse three standard extensions of the basic Dialectica interpretation of intuitionistic logic, namely, extensions to classical logic (Section 5.1), arithmetic (Section 5.2) and mathematical analysis (Section 5.3). The extension to the classical context is normally obtained via a combination of the Dialectica interpretation with one of the possible variants of the *negative translation* [13, 20]. In the case of arithmetic, the

important step is the interpretation of the induction schema. This can be done, as shown by Gödel [14], via a form of *primitive recursion on all finite types*. Finally, Spector [32] solved the problem of the Dialectica interpretation of comprehension via a form of recursion on well-founded trees, known as *bar recursion*. In the following we analyse the role of contraction in each of these extensions.

### 5.1 Extension 1: Classical logic and negative translations

The Dialectica interpretation does not provide a direct interpretation of classical logic, since one cannot in general find witnesses for the interpretation of the *double negation elimination*

(DNE)  $\neg\neg A \rightarrow A$ .

What one can do, however, is to bypass any uses of the double negation elimination via some form of negative translation. Negative translations associate formulas  $A$  with classically equivalent formulas  $(A)^N$ . The benefit is that even when  $A$  is only valid classically, the formula  $(A)^N$  will also be valid intuitionistically. Hence, we can define the Dialectica interpretation of a classical formula  $A$  as the Dialectica interpretation of  $(A)^N$ . These two steps can be combined into a single interpretation, as shown by Shoenfield [31, 33].

In order to see the implicit role of contraction in the principle (DNE) we need to move into *classical linear logic*. Let us define linear negation as  $A^\perp := A \multimap \perp$ . Then, the double (linear) negation elimination principle

(DNE<sub>l</sub>)  $(A^\perp)^\perp \multimap A$

has a trivial Dialectica interpretation, since  $|A^\perp|_y^x \equiv (|A|_x^y)^\perp$ . This indicates that “negation” is not the culprit in (DNE). The difficulty in interpreting classical logic comes from the fact that intuitionistic negation  $\neg A$  ( $\equiv !A \multimap \perp$ ) is weaker than linear negation  $A^\perp$  ( $\equiv A \multimap \perp$ ). In the context of linear logic, the principle corresponding to (DNE) is  $(!(A^\perp)^\perp)^\perp \rightarrow A$ . If we abbreviate  $(!(A^\perp)^\perp)^\perp$  as  $?A$ , we can write this as

(CEE)  $?!A \multimap A$

which we will refer to as the *coupled exponentials elimination*. Similarly to the negative translation of classical into intuitionistic logic, Girard's  $?!$ -translation is precisely the tool one needs to avoid any uses of this principle, and translate a (CEE)-proof of  $A$  into a proof of an equivalent theorem which does not use (CEE).

### 5.2 Extension 2: Arithmetic and Gödel's primitive recursion

As mentioned in the introduction, Gödel's purpose for defining the Dialectica interpretation was actually to interpret full Heyting arithmetic into a quantifier-free calculus (an extension of primitive recursive arithmetic to all finite types). Gödel's interpretation involves reducing induction for *arbitrary* formulas to induction for *quantifier-free* formulas. The interpretation of induction relies on a intricate argument. The problem is that one has to carefully use the pair of witnesses  $s_b, s_f$  for  $A(n) \rightarrow A(n+1)$  in order

to produce a witness for  $A(k)$ , for arbitrary  $k$ . Due to the dependency of  $f$  on  $x$  in (4), in the interpretation of intuitionistic implication, the two witnesses must be used in a subtle way.

As we will see, by moving into the linear logic context this interference of the two flows is restricted to the rule of contraction, which makes the treatment of the rule of induction very smooth. Consider the extension of the system  $\text{LL}_2^\omega$  with the following *linear induction rule*<sup>2</sup>

$$\frac{\vdash A(0) \quad A(n) \vdash A(n+1)}{\vdash A(k)} \text{ (IND)}$$

Let us refer to this system by  $\text{LA}^\omega$  (linear arithmetic). It is easy to see that  $\text{HA}^\omega$  is still embeddable into  $\text{LA}^\omega$  as in Definition 3. The translation of the induction rule of  $\text{HA}^\omega$  is derivable in  $\text{LA}^\omega$  as:

$$\frac{\frac{\vdash (A(0))^* \quad \frac{!(A(n))^* \vdash (A(n+1))^*}{!(A(n))^* \vdash (A(n+1))^*} (!_r)}{\vdash (A(0))^*} (!_r) \quad \frac{!(A(n))^* \vdash (A(n+1))^*}{!(A(n))^* \vdash (A(n+1))^*} (!_r)}{\vdash (A(k))^*} \text{ (IND)} \quad \frac{(A(k))^* \vdash (A(k))^*}{!(A(k))^* \vdash (A(k))^*} (!_l)}{\vdash (A(k))^*} \text{ (cut)}$$

The advantage of working with the rule (IND) of  $\text{LA}^\omega$  is that its Dialectica interpretation is much simpler, since contraction is dealt with separately and the forward and backward flows do not interact. Assuming we have Eloise's winning moves for  $A(n) \multimap A(n+1)$  and  $A(0)$  we can obtain a winning move for  $A(k)$  as:

$$\frac{\frac{\vdash |A(0)|_{\mathbf{y}}^{\mathbf{r}}}{\vdash |A(0)|_{\mathbf{y}}^{\mathbf{R}(s_f, \mathbf{r}, 0)}} \quad \frac{\frac{|A(n)|_{s_b(n, \mathbf{y})}^{\mathbf{x}} \vdash |A(n+1)|_{\mathbf{y}}^{s_f(n, \mathbf{x})}}{|A(n)|_{s_b(k-n, \mathbf{y})}^{\mathbf{R}(s_f, \mathbf{r}, n)} \vdash |A(n+1)|_{\mathbf{y}}^{s_f(n, \mathbf{R}(s_f, \mathbf{r}, n))}} (1)}{\frac{|A(n)|_{s_b(k-n, \mathbf{R}(\tilde{s}_b, \mathbf{y}, k-n-1))}^{\mathbf{R}(s_f, \mathbf{r}, n)} \vdash |A(n+1)|_{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-n-1)}^{s_f(n, \mathbf{R}(s_f, \mathbf{r}, n))}} (2)}{\frac{\vdash |A(0)|_{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-0)}^{\mathbf{R}(s_f, \mathbf{r}, 0)} \quad \frac{|A(n)|_{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-n)}^{\mathbf{R}(s_f, \mathbf{r}, n)} \vdash |A(n+1)|_{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-(n+1))}^{\mathbf{R}(s_f, \mathbf{r}, n+1)}}{|A(n)|_{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-n)}^{\mathbf{R}(s_f, \mathbf{r}, n)} \vdash |A(n+1)|_{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-(n+1))}^{\mathbf{R}(s_f, \mathbf{r}, n+1)}}} (IND)}{\vdash |A(k)|_{\mathbf{y}}^{\mathbf{R}(s_f, \mathbf{r}, k)}} (3)}$$

where  $\tilde{s}_b(k-n, \mathbf{y}) = s_b(n, \mathbf{y})$ ,  $\mathbf{R}$  is Gödel's primitive recursor (cf. [34]), and the following substitutions have been used: (1)  $\frac{\mathbf{R}(s_f, \mathbf{r}, n)}{\mathbf{x}}$ , (2)  $\frac{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-n-1)}{\mathbf{y}}$  and (3)  $\frac{\mathbf{R}(\tilde{s}_b, \mathbf{y}, k-0)}{\mathbf{y}}$ . This shows that we can interpret  $\text{LA}^\omega$  into  $\text{qflLA}^\omega$ . The complexity of interpreting  $\text{HA}^\omega$  is then pushed from the Dialectica interpretation into the embedding of  $\text{HA}^\omega$  into  $\text{LA}^\omega$ .

### 5.3 Extension 3: Mathematical analysis and bar recursion

The most surprising extension of the Dialectica interpretation is due to Spector [32] and covers the whole of classical analysis. As we have seen above, classical logic can be avoided via the use of negative translations, and arbitrary induction can be interpreted via quantifier-free induction and primitive recursion in all finite types. The next step was to interpret the whole of classical analysis by giving an interpretation to comprehension

<sup>2</sup> See the conclusion of Shirahata [30] for a discussion about the different ways of extending linear linear logic with mathematical induction.



$$(CA) \exists f \forall n^i (f(n) = 0 \leftrightarrow A(n))$$

which in the presence of classical logic can be reduced to countable choice

$$(AC_c) \forall n^i \exists x^\tau A(n, x) \rightarrow \exists f \forall n A(n, fn).$$

Spector's bar recursion [21, 32] comes up when one tries to find witnesses for the Dialectica interpretation of the negative translation of  $(AC_c)$  (see also [2, 3]). Given that (in an intuitionistic context) the countable choice itself has a trivial interpretation, interpreting its negative translation boils down to interpreting the double negation shift

$$(DNS) \forall n^i \neg\neg A \rightarrow \neg\neg \forall n A.$$

Let us move briefly to the context of classical linear logic, where we also have the modality  $?A$ , which is the dual of  $!A$ . The modality  $?A$  can be given an interpretation dual to that of  $!A$ , i.e.

$$|?A|_y^f := ?|A|_y^{fy}.$$

As noticed by Girard (cf. Section 5.1), double negations are to intuitionistic logic as  $?!$  is to linear logic. Therefore, the linear logic counterpart of (DNS) is

$$(DNS_l) !\forall n ?!A(n) \multimap ?!\forall n A(n).$$

It is not surprising that in trying to produce a winning move for Eloise for the game  $(DNS_l)$  one is quickly led to solve precisely the same set of equations Spector solved in the intuitionistic context. Our recent analysis of Spector's bar recursion [21] suggests that an interesting finitary version of  $(DNS_l)$  is  $!(?!A \wedge ?!B) \multimap ?!(A \wedge B)$ , which is indeed provable in classical linear logic, but requires extremely tricky solutions as the numbers of formulas in the conjunction grows (see also [10]).

## 6 Variants of Gödel's Dialectica Interpretation

Moving into the linear logic context also helps us to understand the subtle differences between the variants of the Dialectica interpretation. In this section we briefly discuss three of these variants which do not require decidability of quantifier-free formulas, and we sketch the different ways in which the problem of contraction is dealt with. The three variants we consider are: the Diller-Nahm variant [6], the bounded functional interpretation [8, 9] and Kohlenbach's monotone variant of Gödel's Dialectica interpretation [15, 16].

Regarding the *formula interpretation* (i.e., the interpretation of formulas as games) the monotone interpretation is the same as the usual interpretation, whereas the Diller-Nahm and bounded interpretations are less strict with Abelard as he only has to produce a "bound" on his move (and not the precise move). In the case of the Diller-Nahm variant the bound is a finite set of candidate moves, and Abelard wins if any of these is a winning move. As for the bounded functional interpretation, the bound is a majorant

(in the sense of Bezem [4]). In this case, Abelard wins if any move below the bound is a winning move:

	Diller-Nahm	Bounded	Monotone
$! A _f^x : \equiv$	$!\forall y \in fx  A _y^x$	$!\forall y \leq^* fx  A _y^x$	$! A _{fx}^x$

When it comes to the *proof interpretation* (i.e., extraction of Eloise’s winning move from the proof of  $A$ ) then the Diller-Nahm and bounded interpretations are more strict and ask for a precise witness  $a$ , while in the monotone interpretation  $a$  is only required to be a majorant for some witness  $x$ :

	Diller-Nahm	Bounded	Monotone
if $\vdash A$ then	$\vdash \forall y  A _y^a$	$\vdash \forall y  A _y^a$	$\vdash \exists x \leq^* a \forall y  A _y^x$

The need for the decidability of quantifier-free formulas in Gödel’s original interpretation comes from the need to decide which of two candidate witnesses is indeed a valid witness. There are basically two ways of circumventing the need for making such a choice. Either one postpones the choice and simply collects the witnesses into a set (either finite or majorizable set) or one requires the choice to be made but then at the end bounds the choice functions so that it becomes unnecessary. For more details on the different variants of the Dialectica interpretation see [22–24].

**Acknowledgements.** The Dialectica interpretation of linear logic presented here is work of de Paiva [26]. The relation to Gödel’s original Dialectica interpretation is due to Shirahata [30]. Many thanks to Thomas Strahm for coordinating the publication of this special issue, and for the invitation to contribute. The author gratefully acknowledges support of the Royal Society under grant 516002.K501/RH/kk.

## References

1. J. Avigad and S. Feferman. Gödel’s functional (“Dialectica”) interpretation. In S. R. Buss, editor, *Handbook of proof theory*, volume 137 of *Studies in Logic and the Foundations of Mathematics*, pages 337–405. North Holland, Amsterdam, 1998.
2. U. Berger and P. Oliva. Modified bar recursion and classical dependent choice. *Lecture Notes in Logic*, 20:89–107, 2005.
3. U. Berger and P. Oliva. Modified bar recursion. *Mathematical Structures in Computer Science*, 16:163–183, 2006.
4. M. Bezem. Strongly majorizable functionals of finite type: a model for bar recursion containing discontinuous functionals. *The Journal of Symbolic Logic*, 50:652–660, 1985.
5. E. Bishop. Mathematics as a numerical language. In A. Kino, J. Myhill, and R. E. Vesley, editors, *Intuitionism and Proof Theory*, pages 53–71. North Holland, Amsterdam, 1970.
6. J. Diller and W. Nahm. Eine Variant zur Dialectica-interpretation der Heyting Arithmetik endlicher Typen. *Arch. Math. Logik Grundlagenforsch.*, 16:49–66, 1974.
7. S. Feferman. Gödel’s Dialectica interpretation and its two-way stretch. In G. Gottlob et al., editor, *Computational Logic and Proof Theory*, volume 713 of *Lecture Notes in Computer Science*, pages 23–40. Springer, Berlin, 1993.

8. F. Ferreira and P. Oliva. Bounded functional interpretation. *Annals of Pure and Applied Logic*, 135:73–112, 2005.
9. F. Ferreira and P. Oliva. Bounded functional interpretation in feasible analysis. *Annals of Pure and Applied Logic*, 145:115–129, 2007.
10. P. Gerhardy. Functional interpretation and modified realizability interpretation of the double-negation shift. In A. Beckmann, U. Berger, B. Löwe, and J. V. Tucker, editors, *Logical Approaches to Computational Barriers, CiE 2006*, pages 109–118. Report CSR 7-2006, Swansea, 2006.
11. J.-Y. Girard. Linear logic. *Theoretical Computer Science*, 50(1):1–102, 1987.
12. J.-Y. Girard. Towards a geometry of interaction. *Contemporary Mathematics*, 92, 1989.
13. K. Gödel. Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines Mathematischen Kolloquiums*, 4:34–38, 1933.
14. K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12:280–287, 1958.
15. U. Kohlenbach. Effective bounds from ineffective proofs in analysis: an application of functional interpretation and majorization. *The Journal of Symbolic Logic*, 57:1239–1273, 1992.
16. U. Kohlenbach. Analysing proofs in Analysis. In W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, editors, *Logic: from Foundations to Applications*, pages 225–260. Oxford University Press, 1996.
17. U. Kohlenbach and P. Oliva. Proof mining: a systematic way of analysing proofs in mathematics. *Proceedings of the Steklov Institute of Mathematics*, 242:136–164, 2003.
18. U. Kohlenbach and P. Oliva. Proof mining in  $L_1$ -approximation. *Annals of Pure and Applied Logic*, 121:1–38, 2003.
19. G. Kreisel. Interpretation of analysis by means of constructive functionals of finite types. In A. Heyting, editor, *Constructivity in Mathematics*, pages 101–128. North Holland, Amsterdam, 1959.
20. S. Kuroda. Intuitionistische Untersuchungen der formalistischen Logik. *Nagoya Mathematical Journal*, 3:35–47, 1951.
21. P. Oliva. Understanding and using Spector's bar recursive interpretation of classical analysis. In A. Beckmann, U. Berger, B. Löwe, and J. V. Tucker, editors, *Proceedings of CiE'2006, LNCS 3988*, pages 423–234. Springer, 2006.
22. P. Oliva. Unifying functional interpretations. *Notre Dame Journal of Formal Logic*, 47(2):263–290, 2006.
23. P. Oliva. Computational interpretations of classical linear logic. In *Proceedings of WoLLIC'07, LNCS 4576*, pages 285–296. Springer, 2007.
24. P. Oliva. Modified realizability interpretation of classical linear logic. In *Proc. of the Eighteenth Annual IEEE Symposium on Logic in Computer Science LICS'07*. IEEE Press, 2007.
25. V. C. V. de Paiva. The Dialectica categories. In J. W. Gray and A. Scedrov, editors, *Proc. of Categories in Computer Science and Logic, Boulder, CO, 1987*, pages 47–62. Contemporary Mathematics, vol 92, American Mathematical Society, 1989.
26. V. C. V. de Paiva. A Dialectica-like model of linear logic. In D. Pitt, D. Rydeheard, P. Dybjer, A. Pitts, and A. Poigné, editors, *Category Theory and Computer Science, Manchester, UK*, pages 341–356. Springer-Verlag LNCS 389, 1989.
27. V. C. V. de Paiva. Dialectica and Chu constructions: Cousins? *Theory and Applications of Categories*, 17(7):127–152, 2007.
28. H. Schellinx. Some syntactical observations on linear logic. *Journal of Logic and Computation*, 1(4):537–559, 1991.
29. H. Schwichtenberg. Dialectica interpretation of well-founded induction. To appear: *Mathematical Logic Quarterly*.
30. M. Shirahata. The Dialectica interpretation of first-order classical linear logic. *Theory and Applications of Categories*, 17(4):49–79, 2006.

31. J. R. Shoenfield. *Mathematical Logic*. Addison-Wesley Publishing Company, 1967.
32. C. Spector. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles in current intuitionistic mathematics. In F. D. E. Dekker, editor, *Recursive Function Theory: Proc. Symposia in Pure Mathematics*, volume 5, pages 1–27. American Mathematical Society, Providence, Rhode Island, 1962.
33. T. Streicher and U. Kohlenbach. Shoenfield is Gödel after Krivine. *Mathematical Logic Quarterly*, 53:176–179, 2007.
34. A. S. Troelstra. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973.