Defining Formal Explanation in Classical Logic by Substructural Derivability*

Francesco A. Genco and Francesca Poggiolesi

IHPST, Université Paris 1 Panthéon-Sorbonne and CNRS
13, Rue du Four, 75006 Paris, France

Abstract. Precisely framing a formal notion of explanation is a hard problem of great relevance for several areas of scientific investigation such as computer science, philosophy and mathematics. We study a notion of formal explanation according to which an explanation of a formula F must contain all and only the true formulae that concur in determining the truth of F. Even though this notion of formal explanation is defined by reference to derivability in classical logic, the relation that holds between the explained formula and the formulae explaining it has a distinct substructural flavour, due to the fact that no redundancy is admitted among the explaining formulae. We formalise this intuition and prove that this notion of formal explanation is essentially connected, in a very specific sense, to derivability in a substructural calculus.

Keywords: Formal explanation · Substructural logics · Proof theory.

1 Introduction

Precisely framing a notion of explanation in formal contexts is a hard problem of great relevance for several areas of scientific investigation, as, for instance, computer science [12,10], philosophy [1,3], mathematics [19,9], and the natural sciences [8,17]. A research line essentially concerned with the formal definition and study of explanation relations is the one investigating the relation of *logical grounding*. According to this approach, a multiset of formulae Δ formally explains a complex formula F if Δ is a logical ground of F, that is, if F is true in virtue of the truth of the elements of Δ. Several endeavours along these lines have appeared in the contemporary literature, see, for instance, [18,5,16,4,13,14].

According to the strictest, and hence most informative, of these technical notions of formal explanation [13,14], a formal explanation of a true statement F must mention all and only the true statements that concur in determining the truth of F. This characterisation is captured by two formal conditions on the multiset Δ of explaining formulae. First, truth of the formulae in Δ must determine the truth of F. Second, the multiset Δ of explaining formulae must be a maximal multiset of, possibly negated, disjoint occurrences of subformulae of the explained formula F. Even though this notion of formal explanation is

* Supported by project IBS (ANR-18-CE27-0012-01).

defined by reference to derivability in classical logic [13,15], the relation that holds between the explained formula and the multiset of formulae explaining it has a distinct substructural flavour, as discussed in [15]. In the present work we present a study of the formal connection between derivability in substructural logics and this notion of formal explanation.

Intuitively, the complexity constraints required to define this notion—which now on we will simply call *formal explanation*—impose a strict correspondence between the formulae occurring in the explanation and the syntactical parts of the formula explained. A formal explanation, indeed, is supposed to mention exactly those parts of the explained formula in virtue of which it is true, and no redundancy is admitted. Since the substructural logics in which the rules of weakening and contraction are not admissible can be seen as resource-aware reasoning systems in which the usage of redundant or duplicated hypotheses is limited, this feature of explanation raises the question whether there exists a substructural logic for which it is possible to spell out a rigorous connection with the relation of formal explanation. In the present work, we positively answer this question by defining a suitable substructural calculus and formally proving its connection with formal explanation. In order to do so, we define the substructural calculus SL which can be either seen as a fragment of FLe (Full Lambek Calculus with the Exchange rule) [6], or as a fragment of Linear Logic [7] extended by an axiom for the lattice bottom $\bot$.[1] We then show that the relation of formal explanation is essentially connected, in a very specific and formal sense, to derivability in SL.

The article is structured as follows. In Section 2, we define the notion of formal explanation in the traditional way, that is, by employing derivability classical logic. In Section 3, we present the substructural calculus SL that we will employ to show the relationship between formal explanation and substructural derivability. In Section 4, we show that formal explanation can be characterised by employing derivability conditions formulated in the calculus SL. In Section 5, we conclude with some remarks and a discussion of future work directions.

2 Formal Explanation

We present now the relation of formal explanation for classical logic formulae as introduced in [13]. The language that we will employ for classical logic is the following:

$$\varphi ::= \psi \mid \varphi \land \varphi \mid \varphi \lor \varphi \mid \varphi \to \varphi \mid \neg\varphi$$

$$\psi ::= p_1 \mid p_2 \mid p_3 \ldots$$

where $p_1, p_2, p_3 \ldots$ is a list of all the propositional variables of the language. In the following, we will employ capital Latin letters as metavariables for formulae and capital Greek letters as metavariables for multisets of formulae.

[1] The symbol $\bot$ is used in [7] for a weaker falsity constant which is often denoted by 0 in other works, see [11, p. 42] for a comparison table of different notations used for linear logic. In SL, no constant corresponds to Girard's constant 0.

In order to introduce the relation of formal explanation, we have to define some technical notions. We start from the converse of a formula.

Definition 1. *For any natural number n, we indicate by $\neg^n A$ the formula $\neg \dots \neg A$ where A does not have $\neg$ as outermost connective and is preceded by a sequence of n negations $\neg$.*

For any formula A, the notation A^* represents the converse of A:

- *if $A = \neg^{2n} B$ then $A^* = \neg A$*
- *if $A = \neg^{2n+1} B$ then $A^* = \neg^{2n} B$*

For any multiset of formulae Γ, we denote by Γ^* the multiset $\{A^* \mid A \in \Gamma\}$.

Notice in particular that, for any formula A, $(A^*)^* = A$ since $((\neg^{2n} B)^*)^* = (\neg^{2n+1} B)^* = \neg^{2n} B$ and $((\neg^{2n+1} B)^*)^* = (\neg^{2n} B)^* = \neg^{2n+1} B$.

We can now define the notion of explanatory subformulae of a formula (exp-subformulae, for short). Intuitively, a set of exp-subformulae of a formula A contains either a positive occurrence or a negative occurrence (formalised by the notion of converse) of each maximal proper subformula of A the truth of which determines the truth of A.[2]

Definition 2. *The set of sets of exp-subformulae $es(A)$ of a classical formula A is defined as follows:*

- $es(B \wedge C) = es(B \vee C) = es(B \rightarrow C) = es(\neg(B \wedge C)) = es(\neg(B \vee C)) = es(\neg(B \rightarrow C)) = \{\{B, C\}, \{B, C^*\}, \{B^*, C\}, \{B^*, C^*\}\}$
- $es(p) = es(\neg p) = \emptyset$
- $es(\neg\neg B) = \{\{B\}, \{B^*\}\}$

We can finally define our notion of formal explanation for classical logic formulae. In order to do so, we first introduce the relation of *immediate formal explanation*, an instance of which intuitively corresponds to a basic step of explanation which displays the immediate reasons why a formula is true. Then, we introduce the relation of *mediate formal explanation*. This relation generalises that of immediate formal explanation and captures the idea that the process of explaining a statement can be iterated in order to find simpler and simpler reasons for our original statement.

Definition 3. *For any consistent multiset $\{C\} \cup \Gamma$ where C is a formula of classical logic and Γ is a multiset of formulae of classical logic, we say that, under the robust condition C (which might not occur), Γ is an immediate formal explanation of A, in symbols $[C]\Gamma \Vdash A$, if and only if:*

---

[2] For the philosophical motivation behind the definition of exp-subformulae, see the notion of *g-subformulae* in [13]. Notice also that the notion of g-subformula is closed under commutativity and associativity of conjunction and disjunction. In the present work, for the sake of simplicity, we omit this requirement. This, nevertheless, does not imply a loss of generality since also the corresponding substructural connectives $\otimes$ and $\sqcup$ of SL are commutative and associative.

– $\Gamma \vdash A$ *(positive derivability),*
– $C, \Gamma^* \vdash A^*$ *(negative derivability),*
– $\{C\} \cup \Gamma$ *is a set of exp-subformulae of A, see Definition 2.*

where $\vdash$ *indicates derivability any system for classical logic.*

The notion of mediate formal explanation can be defined as follows via the notion of immediate formal explanation.

Definition 4. *For any consistent multiset of classical formulae* $\Gamma \cup \Delta$ *and classical formula* A, $[\Gamma]\Delta \Vdash^m A$ *if, and only if, one of the following conditions is satisfied:*

– $[\Gamma]\Delta \Vdash A$
– $[G]\Delta' \Vdash B$ *and* $[\Gamma']\Delta'', B \Vdash^m A$ *where*
  - $\Delta = \Delta' \cup \Delta''$
  - $\Gamma = \Gamma' \cup \{G\}$.

As shown in [15, pp. 14–16], positive and negative derivability characterize mediate formal explanation as well.

Proposition 1. *For any classical formula* A *and pair of multisets of classical formulae* Γ *and* Δ*, we have that if* $[\Gamma]\Delta \Vdash^m A$ *holds, then* $\Delta \vdash A$ *and* $\Gamma, \Delta^* \vdash A^*$.

Given Definition 3, it is possible to list all potential immediate formal explanations for the complex formulae of classical logic.[3] A complete enumeration of all valid schemata of explanations of this kind is the following:

$A, B \Vdash A \wedge B$		
$A, B \Vdash A \vee B$	$[A^*]B \Vdash A \vee B$	$[B^*]A \Vdash A \vee B$
$A^*, B \Vdash A \to B$	$[A]B \Vdash A \to B$	$[B^*]A^* \Vdash A \to B$
$A \Vdash \neg\neg A$		
$A^*, B^* \Vdash \neg(A \wedge B)$	$[A]B^* \Vdash \neg(A \wedge B)$	$[B]A^* \Vdash \neg(A \wedge B)$
$A^*, B^* \Vdash \neg(A \vee B)$		
$A, B^* \Vdash \neg(A \to B)$		

3 The Substructural Calculus

We introduce now the substructural calculus SL which will be used to show that the derivability conditions defining formal explanations in classical logic can be expressed in a substructural calculus.

The language that we will employ for the substructural calculus SL is the following:

$$\varphi ::= \psi \mid \bot \mid \varphi \otimes \varphi \mid \varphi \sqcup \varphi \mid \varphi \multimap \varphi$$

$$\psi ::= p_1 \mid p_2 \mid p_3 \ldots$$

[3] We talk here of potential explanations because, in order to have an actual explanation, we must also guarantee that all formulae occurring in the explanation are true.

where $p_1, p_2, p_3 \ldots$ is a list of all the propositional variables of the language.

The constant $\bot$ for falsity corresponds to the lattice bottom, see [6, p. 83] and [11, p. 42]; the connective $\otimes$ is the multiplicative conjunction;[4] $\sqcup$ is the additive disjunction;[5] and $\multimap$ is the multiplicative implication.

Rules and axioms of the substructural calculus SL are shown in Table 1.

$$A \Rightarrow A \qquad \Gamma, \bot \Rightarrow C \qquad \frac{\Gamma \Rightarrow A \quad \Delta, A \Rightarrow B}{\Gamma, \Delta \Rightarrow B}\ cut$$

$$\frac{\Gamma \Rightarrow A \quad \Delta \Rightarrow B}{\Gamma, \Delta \Rightarrow A \otimes B}\ \otimes r \qquad \frac{\Gamma, A, B \Rightarrow C}{\Gamma, A \otimes B \Rightarrow C}\ \otimes l$$

$$\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow A \sqcup B}\ \sqcup r \qquad \frac{\Gamma \Rightarrow B}{\Gamma \Rightarrow A \sqcup B}\ \sqcup r \qquad \frac{\Gamma, A \Rightarrow C \quad \Gamma, B \Rightarrow C}{\Gamma, A \sqcup B \Rightarrow C}\ \sqcup l$$

$$\frac{\Gamma, A \Rightarrow B}{\Gamma \Rightarrow A \multimap B}\ \multimap r \qquad \frac{\Gamma \Rightarrow A \quad \Delta, B \Rightarrow C}{\Gamma, \Delta, A \multimap B \Rightarrow C}\ \multimap l$$

Table 1. The substructural calculus SL

The calculus SL can be directly seen as a fragment of FLe (Full Lambek Calculus with the Exchange rule)—see, for instance, [6]. Otherwise, it is possible to consider SL as a fragment of Linear Logic—see, for instance, [11]—to which we add the axiom for the lattice bottom $\bot$.

Definition 5. *We define $\sim A$ as $A \multimap \bot$, and we introduce the following two rules for $\sim$ by simplifying the rules for $\multimap$:*

$$\frac{\Gamma, A \Rightarrow \bot}{\Gamma \Rightarrow \sim A}\ \sim r \qquad \frac{\Gamma \Rightarrow A}{\Gamma, \Delta, \sim A \Rightarrow C}\ \sim l$$

Notice that we omitted a premise of the form $\Delta, \bot \Rightarrow C$ from the $\sim r$ rule since this premise is always an axiom of the calculus.

We extend the definition of converse formula also for the language of SL.

Definition 6. *For any formula A in the language of* SL*, the converse A^* of A is defined according to Definition 1 where instead of the classical negation $\neg$ we use the substructural negation $\sim$.*

[4] The multiplicative conjunction is sometimes also called group-theoretical conjunction [11].

[5] The additive disjunction is sometimes also called lattice-theoretical disjunction [11]. This connective is often denoted by $\vee$, but here we reserve the symbol $\vee$ for the disjunction of classical logic.

4 Formal Explanation by Substructural Proofs

We show that the positive and negative derivability conditions used to define the formal explanation relation for classical logic can be expressed as derivability conditions in SL. In order to do this, we interpret classical formulae as substructural formulae according to the following recursive translation t:

$$\begin{aligned} t(A \wedge B) &= t(A) \otimes t(B) \\ t(A \vee B) &= t(A) \sqcup t(B) \sqcup (t(A) \otimes t(B)) \\ t(A \to B) &= t(\neg A \vee B) = \sim t(A) \sqcup t(B) \sqcup (\sim t(A) \otimes t(B)) \\ t(\neg A) &= \sim t(A) \end{aligned}$$

For any multiset Γ of classical formulae, we denote by $t(\Gamma)$ the multiset $\{t(A) \mid A \in \Gamma\}$ of substructural formulae.

The translations of conjunction and negation are self-explanatory. The case of disjunction requires a clarification. Indeed, it is not enough to translate classical disjunction $\vee$ by the substructural $\sqcup$ because, while the multiset of formulae $\{A, B\}$ is a legitimate immediate formal explanation of $A \vee B$, the substructural sequent $A, B \Rightarrow A \sqcup B$ is not derivable. This is due to the fact that the calculus SL enables us to prove a complex formula only if we have the exact amount of hypotheses required to introduce the outermost connective of the formula, and not one hypothesis more. For $A \sqcup B$, in particular, we have that $A \Rightarrow A \sqcup B$ and $B \Rightarrow A \sqcup B$ are derivable, but $A, B \Rightarrow A \sqcup B$ is not, because, in the latter sequent, one formula between A and B is redundant. In more conceptual terms, if A and B are true, both formulae have to be mentioned in order to provide the complete reason why $A \vee B$ is true, and hence both of them have to be in the formal explanation of $A \vee B$; from a substructural perspective, on the other hand, using both A and B to obtain $A \sqcup B$ is a waste of resources and hence is not allowed. Therefore, in order to faithfully encode the behavior of classical disjunction with respect to formal explanation in our substructural calculus SL, we need to translate classical disjunction by a weakened version of the substructural disjunction $\sqcup$. For similar reasons, we translate a classical implication $A \to B$ via its traditional encoding as the classical formula $\neg A \vee B$.

Before showing that the translation t enables us to formulate positive and negative derivability as derivability in SL, we need to prove two propositions concerning the converse relation.

Proposition 2. *For any formula A and multiset Γ, the sequent $\Gamma, A, A^* \Rightarrow \bot$ is derivable.*

Proof. By Definition 6, we have two cases: (i) $A = \sim^{2n} B$ where B does not have $\sim$ as outermost connective, and (ii) $A = \sim^{2n+1} B$, where B does not have $\sim$ as outermost connective.

If (i), $A^* = \sim A$ and $\Gamma, A, A^* \Rightarrow \bot$ can be derived as shown below on the left. If (ii), $A^* = (\sim^{2n+1} B)^* = \sim^{2n} B$ and $\Gamma, A, A^* \Rightarrow \bot$ can be derived as shown below on the right.

$$\frac{A \Rightarrow A}{\Gamma, A, \sim A \Rightarrow \bot} \sim l \qquad \frac{\sim^{2n} B \Rightarrow \sim^{2n} B}{\Gamma, \sim^{2n+1} B, \sim^{2n} B \Rightarrow \bot} \sim l$$

Proposition 3. *For any formula A, the sequent $A^* \Rightarrow \sim A$ is derivable.*

Proof. By Definition 6, we have two cases: (i) $A = \sim^{2n} B$ where B does not have $\sim$ as outermost connective, and (ii) $A = \sim^{2n+1} B$, where B does not have $\sim$ as outermost connective.

If (i), $A^* = \sim A$ and $A^* \Rightarrow \sim A$ is an axiom and hence derivable.

If (ii), $A^* = (\sim^{2n+1} B)^* = \sim^{2n} B$ and $A^* \Rightarrow \sim A$ can be derived as follows:

$$\cfrac{\cfrac{\sim^{2n} B \Rightarrow \sim^{2n} B}{\sim^{2n} B, \sim^{2n+1} B \Rightarrow \bot} \sim l}{\sim^{2n} B \Rightarrow \sim^{2n+2} B} \sim r$$

Theorem 1. *The positive and negative derivability conditions for formal explanation in classical logic can be expressed as derivability conditions in* SL.

Proof. Since the rules of SL are clearly sound with respect to classical logic, we have that, if the conditions of positive and negative derivability are met with respect to SL through the translation t, then they are also met with respect to classical logic. Which means that if a suitably partitioned multiset of formulae C, Γ, D verifies the complexity conditions of Definition 3 and its translation enjoys positive and negative derivability in SL, then $[C]\Gamma \Vdash D$ holds. Therefore, in order to prove the statement it is enough to show that, if $[C]\Gamma \Vdash D$ holds, then the conditions of positive and negative derivability are met in SL for the translation of the formulae C, D and of the multiset Γ. Hence, we consider all valid instances of the formal explanation relation $[C]\Gamma \Vdash D$, we reason by cases on the logical structure of the classical formula D, and we show in each case that $t(\Gamma) \Rightarrow t(D)$ and $t(C), t(\Gamma)^* \Rightarrow t(D)^*$ are derivable sequents. We only show some interesting cases.

– $A, B \Vdash A \vee B$

Positive derivability:

$$\cfrac{\cfrac{A \Rightarrow A \quad B \Rightarrow B}{A, B \Rightarrow A \otimes B}}{A, B \Rightarrow (A \sqcup B) \sqcup (A \otimes B)} \sqcup r$$

Negative derivability:

$$\cfrac{\cfrac{\cfrac{A^*, B^*, A \Rightarrow \bot \quad A^*, B^*, B \Rightarrow \bot}{A^*, B^*, A \sqcup B \Rightarrow \bot} \sqcup l \quad \cfrac{A^*, B^*, A, B \Rightarrow \bot}{A^*, B^*, A \otimes B \Rightarrow \bot} \otimes l}{A^*, B^*, (A \sqcup B) \sqcup (A \otimes B) \Rightarrow \bot} \sqcup l}{A^*, B^* \Rightarrow \sim ((A \sqcup B) \sqcup (A \otimes B))} \sim r$$

where, by Proposition 2, the sequents $A^*, B^*, A \Rightarrow \bot$ and $A^*, B^*, B \Rightarrow \bot$, and the sequent $A^*, B^*, A, B \Rightarrow \bot$ are derivable.

– $[B^*]A \Vdash A \vee B$

Positive derivability: $$\dfrac{\dfrac{A \Rightarrow A}{A \Rightarrow A \sqcup B}\ \sqcup r}{A \Rightarrow (A \sqcup B) \sqcup (A \otimes B)}\ \sqcup r$$

Negative derivability:

$$\dfrac{\dfrac{\dfrac{B^*, A^*, A \Rightarrow \bot \quad B^*, A^*, B \Rightarrow \bot}{B^*, A^*, A \sqcup B \Rightarrow \bot}\ \sqcup l \quad \dfrac{B^*, A^*, A, B \Rightarrow \bot}{B^*, A^*, A \otimes B \Rightarrow \bot}\ \otimes l}{B^*, A^*, (A \sqcup B) \sqcup (A \otimes B) \Rightarrow \bot}\ \sqcup l}{B^*, A^* \Rightarrow \sim ((A \sqcup B) \sqcup (A \otimes B))}\ \sim r$$

where, by Proposition 2, the sequents $B^*, A^*, A \Rightarrow \bot$ and $B^*, A^*, B \Rightarrow \bot$, and the sequent $B^*, A^*, A, B \Rightarrow \bot$ are all derivable.

– $A \Vdash \neg\neg A$

Positive derivability: $$\dfrac{\dfrac{A \Rightarrow A}{A, \sim A \Rightarrow \bot}\ \sim l}{A \Rightarrow \sim\sim A}\ \sim r$$

Negative derivability: we have to derive the sequent $A^* \Rightarrow (\sim\sim A)^*$. By Definition 6, we have two cases: either $A^* = (\sim^{2n} B)^* = \sim A$ and hence $(\sim\sim A)^* = \sim\sim\sim A$; or $A^* = (\sim^{2n+1} B)^* = \sim^{2n} B$ and hence $(\sim\sim A)^* = \sim A$. If $A^* = \sim A$ and $(\sim\sim A)^* = \sim\sim\sim A$, we have the derivation below on the left. If $A^* = (\sim^{2n+1} B)^* = \sim^{2n} B$ and hence $(\sim\sim A)^* = \sim A = \sim^{2n+2} B$, we have the derivation below on the right.

$$\dfrac{\dfrac{\sim A \Rightarrow \sim A}{\sim A, \sim\sim A \Rightarrow \bot}\ \sim l}{\sim A \Rightarrow \sim\sim\sim A}\ \sim r \qquad \dfrac{\dfrac{\sim^{2n} B \Rightarrow \sim^{2n} B}{\sim^{2n} B, \sim^{2n+1} B \Rightarrow \bot}\ \sim l}{\sim^{2n} B \Rightarrow \sim^{2n+2} B}\ \sim r$$

– $A^*, B^* \Vdash \neg(A \wedge B)$

Positive derivability: derivation below on the left. Negative derivability: derivation below on the right.

$$\dfrac{\dfrac{A^*, B^*, A, B \Rightarrow \bot}{A^*, B^*, A \otimes B \Rightarrow \bot}\ \otimes l}{A^*, B^* \Rightarrow \sim (A \otimes B)}\ \sim r \qquad \dfrac{(A^*)^* \Rightarrow A \quad (B^*)^* \Rightarrow B}{(A^*)^*, (B^*)^* \Rightarrow A \otimes B}\ \otimes r$$

where, by Proposition 2, $A^*, B^*, A, B \Rightarrow \bot$ is a derivable sequent; and, since by Definition 6 we have that $(A^*)^* = A$ and $(B^*)^* = B$, both sequents $(A^*)^* \Rightarrow A$ and $(B^*)^* \Rightarrow B$ are derivable.

– $A^*, B^* \Vdash \neg(A \vee B)$

Positive derivability:

$$\dfrac{\dfrac{\dfrac{A^*, B^*, A \Rightarrow \bot \quad A^*, B^*, B \Rightarrow \bot}{A^*, B^*, A \sqcup B \Rightarrow \bot}\ \sqcup l \quad \dfrac{A^*, B^*, A, B \Rightarrow \bot}{A^*, B^*, A \otimes B \Rightarrow \bot}\ \otimes l}{A^*, B^*, (A \sqcup B) \sqcup (A \otimes B) \Rightarrow \bot}\ \sqcup l}{A^*, B^* \Rightarrow \sim ((A \sqcup B) \sqcup (A \otimes B))}\ \sim r$$

where, by Proposition 2, the sequents $A^*, B^*, A \Rightarrow \bot$ and $A^*, B^*, B \Rightarrow \bot$, and the sequent $A^*, B^*, A, B \Rightarrow \bot$ are derivable.

Negative derivability:
$$\dfrac{\dfrac{(A^*)^* \Rightarrow A \quad (B^*)^* \Rightarrow B}{(A^*)^*, (B^*)^* \Rightarrow A \otimes B}\otimes r}{(A^*)^*, (B^*)^* \Rightarrow (A \sqcup B) \sqcup (A \otimes B)}\sqcup r$$
where, since by Definition 6 we have that $(A^*)^* = A$ and $(B^*)^* = B$, the sequents $(A^*)^* \Rightarrow A$ and $(B^*)^* \Rightarrow B$ are derivable.

The previous proof also indicates that the other obvious choices of substructural connectives would not be suitable to characterise immediate formal explanation. Indeed, for additive conjunction $\sqcap$ and multiplicative disjunction $\oplus$, positive derivability does not hold since the leaves of the following derivations are not derivable:

$$\dfrac{A, B \Rightarrow A \quad A, B \Rightarrow B}{A, B \Rightarrow A \sqcap B} \qquad \dfrac{A \Rightarrow A, B}{A \Rightarrow A \oplus B}$$

Moreover, if we used the weaker falsity constant 0 instead of the lattice bottom $\bot$, Proposition 2 would fail, because $\Gamma, 0 \Rightarrow 0$ with $\Gamma \neq \emptyset$ is not derivable, and thus the rule $\sim l$ would not be strong enough to prove what is required.

We show that the generalisation of immediate formal explanation into mediate formal explanation preserves the embeddability of formal explanations in SL.

Theorem 2. *The construction of mediate formal explanations in classical logic preserves positive and negative derivability in the calculus* SL*.*

Proof. Consider any valid instance $[\Gamma]\Delta \Vdash^m A$ of the mediate formal explanation relation. We show that the substructural sequents $t(\Delta) \Rightarrow t(A)$ and $t(\Gamma), t(\Delta)^* \Rightarrow t(A)^*$ are derivable in SL.

According to Definition 4, the instance $[\Gamma]\Delta \Vdash^m A$ is valid if, and only if, it can be justified by a finite number of instances of the immediate formal explanation relation. The proof is by induction on the number of instances of the immediate formal explanation relation used to justify $[\Gamma]\Delta \Vdash^m A$.

In the base case, $[\Gamma]\Delta \Vdash^m A$ itself is a valid instance $[\Gamma]\Delta \Vdash A$ of immediate formal explanation and, by Theorem 1, we have that $t(\Delta) \Rightarrow t(A)$ and $t(\Gamma), t(\Delta)^* \Rightarrow t(A)^*$ are derivable in SL.

Suppose now that $[\Gamma]\Delta \Vdash^m A$ is justified by $n > 1$ instances of the immediate formal explanation relation. Suppose, moreover, that all instances of mediate formal explanation that are justified by less than n instances of immediate formal explanation correspond to sequents derivable in SL. We show that the SL sequents $t(\Delta) \Rightarrow t(A)$ and $t(\Gamma), t(\Delta)^* \Rightarrow t(A)^*$ are derivable as well.

If $[\Gamma]\Delta \Vdash^m A$ is justified by n instances of the immediate formal explanation relation, then, by Definition 4, we have that there is a valid immediate formal explanation $[G]\Delta' \Vdash B$ and a valid mediate formal explanation $[\Gamma']\Delta'', B \Vdash^m A$ which is justified by less than n instances of the immediate formal explanation relation. Moreover, it holds that $\Delta = \Delta' \cup \Delta''$ and $\Gamma = \Gamma' \cup \{G\}$. By induction hypothesis, we have that the pair of sequents $t(\Delta') \Rightarrow t(B)$ and $t(\Delta''), t(B) \Rightarrow t(A)$, and the pair of sequents $t(G), t(\Delta')^* \Rightarrow t(B)^*$ and $t(\Gamma'), t(\Delta'')^*, t(B)^* \Rightarrow t(A)^*$ are derivable in SL. By applying the cut rule, we can obtain the desired derivation of $t(\Gamma), t(\Delta) \Rightarrow t(A)$:

$$\frac{t(\Delta') \Rightarrow t(B) \quad t(\Delta''), t(B) \Rightarrow t(A)}{t(\Delta'), t(\Delta'') \Rightarrow t(A)}\ cut$$

since $t(\Delta) = t(\Delta') \cup t(\Delta'')$; and the desired derivation of $t(\Gamma), t(\Delta)^* \Rightarrow t(A)^*$:

$$\frac{t(G), t(\Delta')^* \Rightarrow t(B)^* \quad t(\Gamma'), t(\Delta'')^*, t(B)^* \Rightarrow t(A)^*}{t(G), t(\Gamma'), t(\Delta')^*, t(\Delta'')^* \Rightarrow t(A)^*}\ cut$$

since $t(\Delta)^* = t(\Delta')^* \cup t(\Delta'')^*$ and $t(\Gamma) = t(\Gamma') \cup t(G)$.

Since formal explanations are often associated with analytic proofs [16], we also prove that formal explanations in classical logic can always be represented as analytic SL derivations.

Corollary 1. *Formal explanations in classical logic can be represented as cut-free* SL *derivations.*

Proof. Since cut-elimination holds for SL, see, for instance, [2, p. 213], we have that for each SL derivation used in Theorem 2, there is a cut-free SL derivation with the same end-sequent.

5 Conclusions

We have fully characterised the notion of immediate formal explanation for classical logic by the substructural calculus SL, which is both a fragment of FLe (Full Lambek Calculus with the Exchange rule) and a fragment of Linear Logic extended by an axiom for the lattice bottom. Moreover, we have proved that the notion of mediate formal explanation—that is, the transitive closure of immediate formal explanation—preserves all relevant derivability conditions in SL without the *cut* rule. Thus, we have showed that formal explanation, in general, can be fully characterised by complexity constraints and by analytic derivability in SL, without any reference to derivability in classical logic. This notion of explanation, as a consequence, does not only have a substructural flavour, but is actually rigorously connected in an essential way to substructural derivability.

A further question remains open, though. Indeed, we conjecture that instead of defining mediate formal explanation as the transitive closure of immediate formal explanation, it should be possible to define it in a direct way by derivability conditions formulated in SL and suitably adapted complexity conditions. In order to prove this, since we already proved that mediate formal explanation preserves positive and negative derivability in *cut*-free SL, it is enough to prove that, if positive and negative derivability in SL hold for the translation of a suitable multiset of formulae Γ, Δ, A, then $[\Gamma]\Delta \Vdash^m A$ holds as well. Even though we conjecture that this is true, a formal proof of the result seems to require a rather lengthy argument. We leave it, therefore, for future work.

References

1. Barnes, J. (ed.): The Complete Works of Aristotle. Princeton University Press (1984)
2. Belardinelli, F., Jipsen, P., Ono, H.: Algebraic aspects of cut elimination. Studia Logica: An International Journal for Symbolic Logic **77**(2), 209–240 (2004)
3. Bolzano, B.: Theory of science. Oxford: Oxford University Press (2014), translated by Rolf George and Paul Rusnok
4. Correia, F.: Logical grounds. The review of symbolic logic **7**(1), 31–59 (2014)
5. Fine, K.: Guide to Ground, pp. 37–80. Cambridge University Press (2012)
6. Galatos, N., Jipsen, P., Kowalski, T., Ono, H.: Residuated lattices: an algebraic glimpse at substructural logics. Elsevier (2007)
7. Girard, J.Y.: Linear logic. Theoretical computer science **50**(1), 1–101 (1987)
8. Hempel, C.G., Oppenheim, P.: Studies in the logic of explanation. Philosophy of science **15**(2), 135–175 (1948)
9. Mancosu, P.: Mathematical explanation: Problems and prospects. Topoi **20**(1), 97–117 (2001)
10. Miller, T.: Explanation in artificial intelligence: Insights from the social sciences. Artificial intelligence **267**, 1–38 (2019)
11. Paoli, F.: Substructural logics: a primer, vol. 13. Springer Science & Business Media (2013)
12. Pearl, J., Mackenzie, D.: The book of why: the new science of cause and effect. Basic books (2018)
13. Poggiolesi, F.: On defining the notion of complete and immediate formal grounding. Synthese **193**, 3147–3167 (2016)
14. Poggiolesi, F.: On constructing a logic for the notion of complete and immediate formal grounding. Synthese **195**, 1231–1254 (2018)
15. Poggiolesi, F.: A proof-based framework for several types of grounding. Logique et Analyse (2020)
16. Rumberg, A.: Bolzano's concept of grounding (abfolge) against the background of normal proofs. Review of Symbolic Logic **6**(3), 424–459 (2013)
17. Salmon, W.C.: Four decades of scientific explanation. University of Pittsburgh press (2006)
18. Schnieder, B.: A logic for because. Review of Symbolic Logic **4**(3), 445–465 (2011)
19. Steiner, M.: Mathematical explanation. Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition **34**(2), 135–151 (1978)