Math. Struct. in Comp. Science (2009), *vol.* 19, *pp.* 1029–1059. © Cambridge University Press 2009
doi:10.1017/S0960129509990089 First published online 24 August 2009

The algebraic lambda calculus

LIONEL VAUX[†]

Laboratoire de Mathématiques de l'Université de Savoie, UFR SFA, Campus Scientifique, 73376 Le Bourget-du-Lac Cedex, France
E-mail: lionel.vaux@univ-savoie.fr

Received 2 May 2008; revised 23 May 2009

We introduce an extension of the pure lambda calculus by endowing the set of terms with the structure of a vector space, or, more generally, of a module, over a fixed set of scalars. Moreover, terms are subject to identities similar to the usual pointwise definition of linear combinations of functions with values in a vector space. We then study a natural extension of beta reduction in this setting: we prove it is confluent, then discuss consistency and conservativity over the ordinary lambda calculus. We also provide normalisation results for a simple type system.

1. Introduction

Preliminary definitions and notation

Recall that a *rig* (or 'semiring with zero and unity') is the same thing as a unital ring, without the condition that every element admits an additive inverse. Let $\mathsf{R} = (\mathsf{R}, +, 0, \times, 1)$ be a rig: $(\mathsf{R}, +, 0)$ is a commutative monoid, $(\mathsf{R}, \times, 1)$ is a monoid, $\times$ is distributive over $+$ and 0 is absorbing for $\times$. We write $\mathsf{R}^\bullet$ for $\mathsf{R} \setminus \{0\}$. We use letters a, b, c to denote the elements of R, and say that R is *positive* if for all $a, b \in \mathsf{R}$, we have $a + b = 0$ implies $a = 0$ and $b = 0$. An example of a positive rig is $\mathbf{N}$, the set of natural numbers, with the usual operations.

A module over a rig R, or R-*module*, is defined in the same way as a unital module over a ring, again without the condition that every element admits an additive inverse. For any set $\mathscr{X}$, the set of formal finite linear combinations of elements of $\mathscr{X}$ with coefficients in R is the free R-module over $\mathscr{X}$, which we denote by $\mathsf{R}\langle\mathscr{X}\rangle$.

Linearity in the λ-calculus

Girard's linear logic (Girard 1987), by decomposing intuitionistic implication, gave prominence to the computational concept of linearity, while relating it to the usual algebraic notion. A program is said to be linear if it uses its argument exactly once. This vague idea can be made more precise by defining which subterms of a term u are in *linear position* in u:

— in a term that is only a variable x, that occurrence of the variable is in linear position;

[†] This work has been partially funded by the French ANR projet blanc *Curry Howard pour la Concurrence* CHOCO ANR-07-BLAN-0324.

— in an abstraction $u = \lambda x\, s$, the subterms in linear position in u are those of the abstracted subterm s, and u itself;
— in an application $u = (s)\, t$, the subterms in linear position in u are those of the function subterm s, and u itself.

In particular, application is linear in the function but not in the argument. This can be related to head reduction and memory management: those subterms that are in linear position are evaluated exactly once in the head reduction, they are neither copied nor discarded.

Algebraic linearity is generally thought of as commutation with sums. It is well known that the space of all functions from some set to some fixed R-module is itself an R-module, with operations on functions defined pointwise: for instance, the sum of two functions is defined by $(f + g)(x) = f(x) + g(x)$. In Ehrhard (2001; 2005), Ehrhard introduced denotational models of linear logic where formulas are interpreted as particular vector spaces or modules, and proofs corresponding to λ-terms are interpreted as analytic functions defined by power series on these spaces: this is the basic idea of Girard's quantitative semantics (Girard 1988). This not only guided the study of differentiation in λ-calculus presented in Ehrhard and Regnier (2003), but also offered serious grounding for endowing the set of terms with the structure of a vector space, or of an R-module, where R is a rig – one can form linear combinations of terms, subject to the following two identities:

$$\lambda x \left(\sum_{i=1}^{n} a_i s_i \right) = \sum_{i=1}^{n} a_i \lambda x\, s_i \tag{1}$$

and

$$\left(\sum_{i=1}^{n} a_i s_i \right) u = \sum_{i=1}^{n} a_i (s_i)\, u \tag{2}$$

for all linear combination $\sum_{i=1}^{n} a_i s_i$ of terms. Here we recover the fact that application is linear in the function and not in the argument, in accordance with the computational notion of linearity.

Reducing linear combinations of λ-terms

Apart from differentiation, an important feature of the calculus of Ehrhard and Regnier (2003) is the way β-reduction is extended to such linear combinations of terms. Among the terms, some are considered simple since they contain no sum in linear position, and thus neither (1) nor (2) applies, so intrinsically they are not sums. These form a basis of the R-module of terms. Reduction $\rightarrow$ is then the least contextual relation such that if s is a simple term,

$$(\lambda x\, s)\, t \rightarrow s\,[t/x], \tag{3}$$

and if $a \in \mathsf{R}^{\bullet}$ is a non-zero scalar,

$$s \rightarrow s' \text{ implies } as + t \rightarrow as' + t. \tag{4}$$

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

Since every ordinary λ-term can be viewed as a simple term, (3) extends the usual β-reduction. The requirement that s is simple in (3) and (4), together with the condition $a \neq 0$ in (4), ensures $\to$ actually reduces something so that reduction is not trivially reflexive.

Although the previous definition might seem contorted, it is technically efficient. For instance, it is particularly well suited for proving confluence using the usual Tait–Martin-Löf technique, which involves introducing a parallel version $\rightrightarrows$ of $\to$ such that $\to \subseteq \rightrightarrows \subseteq \to^*$ and proving that $\rightrightarrows$ enjoys the diamond property. Here $\rightrightarrows$ is reflexive and has the following behaviour on linear combinations of terms:

$$\sum_{i=1}^{n} a_i s_i \rightrightarrows \sum_{i=1}^{n} a_i s_i' \text{ when, for all } i,\ s_i \rightrightarrows s_i' \text{ and } s_i \text{ is simple.} \tag{5}$$

Assuming $s \rightrightarrows s' \rightrightarrows s''$ are simple terms, we have $s + s' \rightrightarrows 2s'$ and $s + s' \rightrightarrows s + s''$. (5) then allows us to close that pair of reductions by $2s' \rightrightarrows s' + s''$ and $s + s'' \rightrightarrows s' + s''$. This would not hold if we had forced the s_i's in (5) to be distinct simple terms – that condition would amount to reducing each element of the base of simple terms in parallel, which, at first, might seem a natural choice.

Collapse

However, Vaux (2007a) proved that the above higher-order rewriting of linear combinations collapses as soon as the rig of scalars admits negative elements since if we have $-1 \in \mathsf{R}$ (so $1 + (-1) = 0$), then for all terms s and t, we have $s \to^* t$. This should not be a surprise, since in that case the system involves both negative numbers and potential infinity through arbitrary fixed points. Indeed, take Θ to be a fixpoint operator of the λ-calculus such that $(\Theta)\, s \to^* (s)\,(\Theta)\, s$ for all λ-term s, and write ∞_s for $(\Theta)\, \lambda x\, (s + x)$. Then $\infty_s \to^* s + \infty_s$, so ∞_s stands for an infinite amount of s. We get

$$s = s + \infty_s - \infty_s + \infty_t - \infty_t \to^* s - s + t = t.$$

Also, if one can consider fractions of scalars, strong normalisability holds only for normal terms, since if we assume $s \to s'$ and that R contains dyadic rationals, then

$$s \;=\; \frac{1}{2}s + \frac{1}{2}s \;\to\; \frac{1}{2}s + \frac{1}{2}s' \;\to\; \frac{1}{4}s + \frac{3}{4}s' \;\to\; \cdots$$

Both these failures indicate that much care is needed when dealing with linear combinations of λ-terms: these make the identity of terms very intricate, much more so than plain α-equivalence, so its interaction with higher-order rewriting becomes tricky. As a result, although the problems with normalisability were well noted in Ehrhard and Regnier (2003), the collapse of reduction in the presence of negative coefficients eluded the authors of that paper. In the present contribution, we give a syntactic framework for the study of linear combinations of terms, which aims to be more rigorous and formal than that developed in Ehrhard and Regnier (2003) or Vaux (2007a). In particular, we devote much care to developing an explicit implementation of the R-module of terms. Also, we do not consider differentiation or classical control operators, and only focus on the algebraic structure of terms and the interaction between coefficients and reduction. We call the resulting system the *algebraic λ-calculus*.

Contributions

In Section 2, we formalise the definition of the R-module of terms, validating identities (1) and (2), and introduce the key notion of canonical forms. We also compare this presentation to that of Ehrhard and Regnier (2003): terms *à la* Ehrhard–Regnier are just canonical forms of terms in our setting. This is an important part of the present work, which we hope sheds new light on the structure of the R-module of terms. In Section 3 we define reduction, using rule (4) in the case of a sum, and discuss conservativity with respect to ordinary β-reduction. Section 4 presents a Curry-style simple type system for the algebraic λ-calculus. We prove that subject reduction holds if and only if the rig of scalars is positive. In Section 5, we discuss necessary conditions for the strong normalisation of typed terms to hold, and then refine these to sufficient conditions and generalise the proof of strong normalisation of the differential λ-calculus given in Ehrhard and Regnier (2003). We conclude by discussing possible other approaches in Section 6.

Previous work

Most of the results of this paper have already been presented in Vaux (2007a) or even Ehrhard and Regnier (2003), though sometimes in a weaker form. In this earlier work, however, the focus was on differentiation, and the presence of linear combinations of terms and their effects on reduction were considered of marginal interest. As we stated before, this may, in particular, explain why some of the problems we concentrate on in this paper were put aside in Ehrhard and Regnier (2003). The material of Sections 2 and 3 was the subject of the RTA'07 conference extended abstract Vaux (2007b). Although a very brief outline of a preliminary version of Section 5 was given in that paper, the normalisation results of the present article are completely new in that they strictly generalise those of Vaux (2007a).

2. Linear combinations of terms

In this section, we introduce the set of terms of the algebraic λ-calculus in several steps. First we give a grammar of terms, on which we define α-equivalence and substitution as in Krivine's (Krivine 1990). Then we define a notion of algebraic equality on these terms: this is given by an equivalence relation $\triangleq$ on terms such that the associated quotient set is an R-module, moreover validating identities (1) and (2). The elements of this quotient set are the objects of the algebraic λ-calculus. We then introduce canonical forms of terms as distinguished elements of $\triangleq$-equivalence classes. We show this construction encompasses the abstract presentation by Ehrhard and Regnier in (Ehrhard and Regnier 2003), based on an increasing sequence of quotients.

2.1. *Raw terms*

Let there be given a denumerable set $\mathscr{V}$ of variables. We use letters among x, y, z to denote variables.

Definition 2.1. The language L^0_R of the *raw terms* of the algebraic λ-calculus over R (denoted by capital letters L, M, N) is given by the following grammar:

$$M, N, \ldots ::= x \mid \lambda x\, M \mid (M)\, N \mid \mathbf{0} \mid aM \mid M + N.$$

Definition 2.2. We define free variables of terms as follows:

— variable x is free in term y if $x = y$;
— variable x is free in $\lambda y\, M$ if $x \neq y$ and x is free in M;
— variable x is free in $(M)\, N$ if x is free in M or in N;
— variable x is free in aM if x is free in M;
— variable x is free in term $M + N$ if x is free in M or in N.

In particular, no variable occurs free in term $\mathbf{0}$. Notice, however, that, by the previous definition, aM might have free variables even if $a = 0$: as far as raw terms are concerned, $0M$ is not the same as $\mathbf{0}$.

From this definition of free variables, we derive α-equivalence (denoted $\sim$) as in Krivine (1990). We will always consider raw terms up-to α-equivalence. More formally, we have the following definition.

Definition 2.3. The set L_R of the raw terms of the algebraic λ-calculus over R is the quotient set $\mathsf{L}^0_\mathsf{R}/\!\sim$.

Again, we derive the definition of substitution following that in Krivine (1990). We write $M\,[N/x]$ for the (capture-avoiding) substitution of N for x in M. More generally, if $x_1, \ldots, x_n$ are distinct variables and $N_1, \ldots, N_n$ are terms, we write

$$M\,[N_1, \ldots, N_n/x_1, \ldots, x_n]$$

for the simultaneous capture avoiding substitution of each N_i for each x_i in M. We obtain the following variants of definitions and properties from Krivine (1990).

Proposition 2.4. For all terms $M, N_1, \ldots, N_n, L_1, \ldots, L_p$ and all distinct variables $x_1, \ldots, x_n$, $y_1, \ldots, y_p$,

$$M\,[N_1, \ldots, N_n/x_1, \ldots, x_n]\,[L_1, \ldots, L_p/y_1, \ldots, y_p] \\ \sim M\,[N'_1, \ldots, N'_n, L_1, \ldots, L_p/x_1, \ldots, x_n, y_1, \ldots, y_p]$$

where $N'_i = N_i\,[L_1, \ldots, L_p/y_1, \ldots, y_p]$.

Definition 2.5. A binary relation r on raw terms is said to be *contextual* if it satisfies the following conditions:

— $x \mathrel{\mathrm{r}} x$;
— $\lambda x\, M \mathrel{\mathrm{r}} \lambda x\, M'$ when $M \mathrel{\mathrm{r}} M'$;
— $(M)\, N \mathrel{\mathrm{r}} (M')\, N'$ when $M \mathrel{\mathrm{r}} M'$ and $N \mathrel{\mathrm{r}} N'$;
— $\mathbf{0} \mathrel{\mathrm{r}} \mathbf{0}$;
— $aM \mathrel{\mathrm{r}} aM'$ when $M \mathrel{\mathrm{r}} M'$;
— $M + N \mathrel{\mathrm{r}} M' + N'$ when $M \mathrel{\mathrm{r}} M'$ and $N \mathrel{\mathrm{r}} N'$.

This notion of a contextual relation is the analogue of a λ-compatible relation in Krivine (1990). In particular, a binary relation r is contextual if and only if it is reflexive and:

— $\lambda x\, M \mathrel{\mathrm{r}} \lambda x\, M'$ when $M \mathrel{\mathrm{r}} M'$;
— $(M)\, N \mathrel{\mathrm{r}} (M')\, N'$ when $M \mathrel{\mathrm{r}} M'$ and $N \mathrel{\mathrm{r}} N'$;
— $aM \mathrel{\mathrm{r}} aM'$ when $M \mathrel{\mathrm{r}} M'$;
— $M + N \mathrel{\mathrm{r}} M' + N'$ when $M \mathrel{\mathrm{r}} M'$ and $N \mathrel{\mathrm{r}} N'$.

Proposition 2.6. If r is a contextual relation, then $M\,[N/x] \mathrel{\mathrm{r}} M\,[N'/x]$ when $N \mathrel{\mathrm{r}} N'$.

Again, this result is only an obvious variant of that given in Krivine (1990).

2.2. *The module of terms*

We introduce the actual algebraic content of the calculus by defining an equivalence relation $\triangleq$ encompassing the usual identities between linear combinations, together with (1) and (2).

Definition 2.7. *Algebraic equality* $\triangleq$ is defined on raw terms as the least contextual equivalence relation such that the following identities hold:

— axioms of a commutative monoid:

$$\mathbf{0} + M \triangleq M \tag{6a}$$
$$(M + N) + L \triangleq M + (N + L) \tag{6b}$$
$$M + N \triangleq N + M \tag{6c}$$

— axioms of a module over rig R:

$$a(M + N) \triangleq aM + aN \tag{7a}$$
$$aM + bM \triangleq (a + b)M \tag{7b}$$
$$a(bM) \triangleq (ab)M \tag{7c}$$
$$1M \triangleq M \tag{7d}$$
$$0M \triangleq \mathbf{0} \tag{7e}$$
$$a\mathbf{0} \triangleq \mathbf{0} \tag{7f}$$

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

— linearity in the λ-calculus:

$$\lambda x\, \mathbf{0} \triangleq \mathbf{0} \tag{8a}$$

$$\lambda x\, (aM) \triangleq a\, (\lambda x\, M) \tag{8b}$$

$$\lambda x\, (M+N) \triangleq \lambda x\, M + \lambda x\, N \tag{8c}$$

$$(\mathbf{0})\, L \triangleq \mathbf{0} \tag{8d}$$

$$(aM)\, L \triangleq a\, ((M)\, L) \tag{8e}$$

$$(M+N)\, L \triangleq (M)\, L + (N)\, L. \tag{8f}$$

We call the elements of $\mathsf{L}_{\mathsf{R}}/\triangleq$, that is, the $\triangleq$-classes of raw terms, *algebraic λ-terms*. If $M \in \mathsf{L}_{\mathsf{R}}$, we write $\underline{M}$ for its $\triangleq$-class.

Notice that identity (7f) could be removed, as it is derived from (7e) and (7c). Identities (8a) through (8c) subsume (1) and identities (8d) through (8f) subsume (2). Then the quotient set $\mathsf{L}_{\mathsf{R}}/\triangleq$ is an R-module validating (1) and (2).

Definition 2.8. For all $M_1, \ldots, M_n \in \mathsf{L}_{\mathsf{R}}$, we write $M_1 + \cdots + M_n$ or even $\sum_{i=1}^{n} M_i$ for the term $M_1 + (\cdots + M_n)$ (or $\mathbf{0}$ if $n = 0$).

One might think of a raw term $M \in \mathsf{L}_{\mathsf{R}}$ as a *writing* of its $\triangleq$-class, which is an element of the R-module $\mathsf{L}_{\mathsf{R}}/\triangleq$. Among raw terms, some should be distinguished as canonical writings. More precisely, we want to make the following statement meaningful: every term $M \in \mathsf{L}_{\mathsf{R}}$ can be uniquely written as $M \triangleq \sum_{i=1}^{n} a_i s_i$ where the s_i's are pairwise distinct *base elements* and the a_i's are non-zero.

A good candidate for such a canonical base is obtained as follows:

— all the identities in groups of equations (6) (7) and (8), except (6c), can be oriented from left to right to form a rewrite system;

— raw terms that are normal in this rewrite system, and are of the shape x, $\lambda x\, M$ or $(M)\, N$, can be considered as base elements (they are not sums);

— every $M \in \mathsf{L}_{\mathsf{R}}$ has a normal form in this system, which can be written as a linear combination of base terms.

Notice, however, that a normal form in this system need not be canonical: consider, for example, $x + y + x$. Of course, the problem is that we left out commutativity: adding (6c) would break the very notion of a normal form. Rewriting up to commutativity, or up to associativity and commutativity, is a notable trend in rewriting theory, with a well-established literature: we will just cite Peterson and Stickel (1981) as an example. Even closer to our subject, Arrighi and Dowek (2005) proposed an associative–commutative rewrite system implementing a computational notion of vector space, which is very close to what we have just outlined.

In the current setting, however, our focus is on specifying the syntax of the algebraic λ-calculus, and we are only interested in the definition of canonical forms and base elements. Hence we will not fully reproduce such a rewrite-theoretic development. Instead, we extend our notion of equality of terms *a minima* so that the order of summands in $\sum_{i=1}^{n} M_i$ no longer matters. As far as syntax is concerned, this is quite benign. Moreover, the reduction of the algebraic λ-calculus, to be defined in Section 3, is introduced as a relation on $\mathsf{L}_\mathsf{R}/\triangleq$: associativity and commutativity will be dissolved in $\triangleq$.

Definition 2.9. *Permutative equality* $\equiv\, \subseteq \mathsf{L}_\mathsf{R} \times \mathsf{L}_\mathsf{R}$ is the least contextual equivalence relation such that $\sum_{i=1}^{n} M_i \equiv \sum_{i=1}^{n} M_{f(i)}$ holds for all $M_1, \ldots, M_n \in \mathsf{L}_\mathsf{R}$ and all permutation f of $\{1, \ldots, n\}$.

Since free variables of a sum do not depend on the order of the summands, $\equiv$ preserves free variables.

Definition 2.10. We write Λ_R for the quotient set $\mathsf{L}_\mathsf{R}/\equiv$, and we call the elements of Λ_R *permutative terms.*

Proposition 2.11. Substitution is well defined on Λ_R. That is, if $M, M' \in \mathsf{L}_\mathsf{R}$ are such that $M \equiv M'$ and for all $i \in \{1, \ldots, n\}$ we have $N_i, N_i' \in \mathsf{L}_\mathsf{R}$ are such that $N_i \equiv N_i'$, then $M[N_1, \ldots, N_n/x_1, \ldots, x_n] \equiv M'[N_1', \ldots, N_n'/x_1, \ldots, x_n]$ for all pairwise distinct variables $x_1, \ldots, x_n$.

Except when stated otherwise, we will use the same notation for a raw term M and its $\equiv$-class, and use them interchangeably. This is harmless in general: the properties we consider are all invariant under $\equiv$ and we define functions on Λ_R by induction on raw terms, compatibility with $\equiv$ being obvious.

Note that algebraic equality already subsumes permutative equality on raw terms, so $\triangleq$ is well defined on Λ_R and $(\mathsf{L}_\mathsf{R}/\triangleq) = (\Lambda_\mathsf{R}/\triangleq)$.

2.3. *Canonical forms*

We can now define canonical forms of terms as particular permutative terms such that every class in $\Lambda_\mathsf{R}/\triangleq$ contains exactly one canonical element.

Definition 2.12. We define the set $\mathsf{C}_\mathsf{R} \subset \Lambda_\mathsf{R}$ of *canonical terms* (denoted by capital letters S, T, U, V, W) and the set $\mathsf{B}_\mathsf{R} \subset \mathsf{C}_\mathsf{R}$ of *base terms* (denoted by small letters s, t, u, v, w) by mutual induction as follows:

- — any variable x is a base term;
- — if $x \in \mathscr{V}$ and s is a base term, then $\lambda x\, s$ is a base term;
- — if s a base term and T is a canonical term, then $(s)\, T$ is a base term;
- — if $a_1, \ldots, a_n \in \mathsf{R}^\bullet$ and $s_1, \ldots, s_n$ are pairwise distinct base terms, then $\sum_{i=1}^{n} a_i s_i$ is a canonical term.

An easy intuition is that for all canonical terms $S, T \in \mathsf{C}_\mathsf{R}$, we have $S \triangleq T$ if and only if $S = T$ (a formal proof of this result is given later as a corollary of Theorem 2.17). Mapping s to the 'singleton' $1s$ defines an injection from base terms into canonical terms.

Definition 2.13. We define the *height* of base terms and canonical terms by mutual induction:

- — $\mathrm{h}(x) = 1$;
- — $\mathrm{h}(\lambda x\, s) = 1 + \mathrm{h}(s)$;

— $\mathrm{h}((s)\,T) = 1 + \max(\mathrm{h}(s), \mathrm{h}(T))$;
— $\mathrm{h}(\sum_{i=1}^{n} a_i s_i) = \max_{1\leqslant i\leqslant n}(\mathrm{h}(s_i))$ (which is 0 if and only if $n = 0$).

Definition 2.14. Let $M = \sum_{i=1}^{n} a_i s_i \in \Lambda_{\mathsf{R}}$ be a linear combination of base terms, which is not necessarily canonical. For any base term s, we use $M_{(s)}$ to denote the scalar $\sum_{1\leqslant i\leqslant n,\, s_i=s} a_i$ (the sum of those a_i's such that $s_i = s$) and call it the *coefficient of s in M*. Then we define $\mathsf{cansum}\,(M) \in \mathsf{C}_{\mathsf{R}}$ by

$$\mathsf{cansum}\,(M) = \sum_{j=1}^{p} M_{(t_j)} t_j$$

where $\{t_1, \ldots, t_p\}$ is the set of those s_i's with a non-zero coefficient in M.

We now define a function mapping terms in Λ_{R} to their canonical forms.

Definition 2.15. Canonisation of terms, $\mathsf{can} : \Lambda_{\mathsf{R}} \longrightarrow \mathsf{C}_{\mathsf{R}}$, is given by:

— $\mathsf{can}\,(x) = 1x$;
— if $\mathsf{can}\,(M) = \sum_{i=1}^{n} a_i s_i$, then $\mathsf{can}\,(\lambda x\, M) = \sum_{i=1}^{n} a_i\,(\lambda x\, s_i)$;
— if $\mathsf{can}\,(M) = \sum_{i=1}^{n} a_i s_i$ and $\mathsf{can}\,(N) = T$, then $\mathsf{can}\,((M)\,N) = \sum_{i=1}^{n} a_i (s_i)\,T$;
— $\mathsf{can}\,(\mathbf{0}) = \mathbf{0}$;
— if $\mathsf{can}\,(M) = \sum_{i=1}^{n} a_i s_i$, then $\mathsf{can}\,(aM) = \mathsf{cansum}\left(\sum_{i=1}^{n} (aa_i) s_i\right)$;
— if $\mathsf{can}\,(M) = \sum_{i=1}^{n} a_i s_i$ and $\mathsf{can}\,(N) = \sum_{i=n+1}^{n+p} a_i s_i$, then

$$\mathsf{can}\,(M + N) = \mathsf{cansum}\left(\sum_{i=1}^{n+p} a_i s_i\right).$$

Notice that in the penultimate case (the definition of $\mathsf{can}\,(aM)$), the only effect of the application of cansum is to prune all the summands $(aa_i)s_i$ such that $aa_i = 0$.

Lemma 2.16. Canonisation enjoys the following properties:

(i) Variables free in $\mathsf{can}\,(M)$ are also free in M. The converse does not hold in general.
(ii) For all base term s, $\mathsf{can}\,(s) = 1s$.
(iii) For all canonical term S, $\mathsf{can}\,(S) = S$.
(iv) For all term $M \in \Lambda_{\mathsf{R}}$, $\mathsf{can}\,(\mathsf{can}\,(M)) = \mathsf{can}\,(M)$.
(v) For all $M, N_1, \ldots, N_n \in \Lambda_{\mathsf{R}}$ and all variables $x_1, \ldots, x_n$ not free in any of these terms, we have

$$\mathsf{can}\left(M\,[N_1, \ldots, N_n/x_1, \ldots, x_n]\right) = \mathsf{can}\left(\mathsf{can}\,(M)\,[\mathsf{can}\,(N_1), \ldots, \mathsf{can}\,(N_n)\,/x_1, \ldots, x_n]\right).$$

Proof. Property (i) is straightforward from the previous definition. Properties (ii) and (iii) are proved by mutual induction on the definitions of base terms and canonical terms. Property (iv) follows from (iii). Property (v) is proved by induction on M, with all inductive steps following directly from the definitions of canonisation and substitution. □

Theorem 2.17. Algebraic equality is equality of canonical forms: for all $M, N \in \Lambda_{\mathsf{R}}$ $M \triangleq N$ if and only if $\mathsf{can}\,(M) = \mathsf{can}\,(N)$.

Proof. For all $M, N \in \Lambda_{\mathsf{R}}$, we write $M \triangleq' N$ if and only if $\mathsf{can}(M) = \mathsf{can}(N)$. It should be clear that $\triangleq'$ is an equivalence relation. It is contextual because the definition of canonisation is by induction on permutative terms. Moreover, it validates equations (6a) through (8f): just apply can to both members of each equation and conclude. By the definition of $\triangleq$, we get $\triangleq \subseteq \triangleq'$. Conversely, one can easily check that $\mathsf{can}(M) \triangleq M$ for all $M \in \Lambda_{\mathsf{R}}$: this is the whole point of the definition of canonisation. Hence, we have the reverse inclusion, *viz.*, if $M \triangleq' N$, then $M \triangleq \mathsf{can}(M) = \mathsf{can}(N) \triangleq N$. □

Corollary 2.18. For all $S, T \in \mathsf{C}_{\mathsf{R}}$, $S \triangleq T$ if and only if $S = T$.

Proof. This a direct consequence of the previous theorem and Lemma 2.16 (iii). □

Corollary 2.19. Substitution is well defined on $\Lambda_{\mathsf{R}}/\triangleq$. That is, if $M, M' \in \Lambda_{\mathsf{R}}$ are such that $M \triangleq M'$ and for all $i \in \{1, \dots, n\}$ we have $N_i, N_i' \in \Lambda_{\mathsf{R}}$ are such that $N_i \triangleq N_i'$, then $M[N_1, \dots, N_n/x_1, \dots, x_n] \triangleq M'[N_1', \dots, N_n'/x_1, \dots, x_n]$ for all pairwise distinct variables $x_1, \dots, x_n$.

Proof. We first apply Theorem 2.17 to the hypotheses and conclusion: we must prove

$$\mathsf{can}\left(M[N_1, \dots, N_n/x_1, \dots, x_n]\right) = \mathsf{can}\left(M'[N_1', \dots, N_n'/x_1, \dots, x_n]\right)$$

knowing that $\mathsf{can}(M) = \mathsf{can}\left(M'\right)$ and, for all $i \in \{1, \dots, n\}$, $\mathsf{can}(N_i) = \mathsf{can}\left(N_i'\right)$. We can then conclude the proof using Lemma 2.16 (v). □

Corollary 2.20. We can define an R-module structure on C_{R} as follows:

$$\begin{aligned} \text{zero:} &\quad \mathbf{0} \in \mathsf{C}_{\mathsf{R}} \\ \text{sum:} &\quad (S, T) \in \mathsf{C}_{\mathsf{R}} \times \mathsf{C}_{\mathsf{R}} \mapsto \mathsf{can}(S + T) \in \mathsf{C}_{\mathsf{R}} \\ \text{scalar multiplication:} &\quad (a, S) \in \mathsf{R} \times \mathsf{C}_{\mathsf{R}} \mapsto \mathsf{can}(aS) \in \mathsf{C}_{\mathsf{R}}. \end{aligned}$$

So can is an isomorphism of R-modules from $\Lambda_{\mathsf{R}}/\triangleq$ to C_{R}.

Proof. By Theorem 2.17, can is well defined on $\Lambda_{\mathsf{R}}/\triangleq$, and is injective. It is surjective by Lemma 2.16 (iii). The R-module structure of C_{R} then follows from that of $\Lambda_{\mathsf{R}}/\triangleq$. □

By this isomorphism, and $\triangleq$ being contextual, the quotient structure of algebraic terms is subsumed by the mutually inductive structure of base terms and canonical terms. If $\mathscr{C}$ is a set of canonical terms, we write $\underline{\mathscr{C}} = \{\underline{S};\ S \in \mathscr{C}\}$, so $(\Lambda_{\mathsf{R}}/\triangleq) = \underline{\mathsf{C}_{\mathsf{R}}}$. When we prove properties on algebraic terms, we can thus use induction on base terms and canonical terms. We then check that the corresponding property on algebraic terms follows through can, which is in general obvious. We will abuse terminology by claiming our proof is by induction on algebraic terms. Also, we will often define functions on $\Lambda_{\mathsf{R}}/\triangleq$ by induction on base terms and canonical terms: the actual function is obtained by composition with can. For instance, we define the height of algebraic terms by $\mathrm{h}(\underline{M}) = \mathrm{h}(\mathsf{can}(\underline{M}))$.

2.4. *Abstract presentation*

Our presentation of the R-module of terms differs from that given in Ehrhard and Regnier (2003) in that we explicitly introduce two distinct levels of syntax: permutative terms on the one hand (Λ_{R}) and algebraic terms on the other ($\Lambda_{\mathsf{R}}/\triangleq$).

One can see the R-module of canonical terms from Corollary 2.20 as a concrete presentation of that adopted by Ehrhard and Regnier, which defines an increasing sequence $(\mathsf{R}\langle\Delta_{\mathsf{R}}(k)\rangle)_{k\geqslant 0}$ of free R-modules generated by simple terms of bounded height.

Definition 2.21. We define the set $\Delta_{\mathsf{R}}(k)$ of *simple terms of height at most k* by induction on k. Let $\Delta_{\mathsf{R}}(0) = \varnothing$. We define the elements of $\Delta_{\mathsf{R}}(k+1)$ from those of $\Delta_{\mathsf{R}}(k)$ by the following clauses:

— if $\sigma \in \Delta_{\mathsf{R}}(k)$, then $\sigma \in \Delta_{\mathsf{R}}(k+1)$;
— if $x \in \mathscr{V}$, then $x \in \Delta_{\mathsf{R}}(k+1)$;
— if $\sigma \in \Delta_{\mathsf{R}}(k)$, then $\lambda x\, \sigma \in \Delta_{\mathsf{R}}(k+1)$;
— if $\sigma \in \Delta_{\mathsf{R}}(k)$ and $\tau \in \mathsf{R}\langle\Delta_{\mathsf{R}}(k)\rangle$, then $(\sigma)\,\tau \in \Delta_{\mathsf{R}}(k+1)$.

Then we define the set of all *simple terms* as $\Delta_{\mathsf{R}} = \bigcup_k \Delta_{\mathsf{R}}(k)$ and the set of *terms* $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle = \bigcup_k \mathsf{R}\langle\Delta_{\mathsf{R}}(k)\rangle$.

Note that although it was not made clear in the original paper, two quotient constructions are interleaved at each height: α-equivalence and the free R-module construction. In our opinion, this makes for a very intricate notion of equality on terms, with the result that the status of prominent and well-established notions in the setting of the ordinary λ-calculus becomes less immediate. For instance, we may ask: What is a free occurrence of a variable in a term? How do we properly define α-conversion on $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$? What are the subterms of a term? Of course, satisfactory answers can be given to these questions: we only claim that the simplicity of the definition is deceptive.

As expected, $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ and $(\Lambda_{\mathsf{R}}/\triangleq)$ are actually the same R-module of algebraic terms. If we define $\mathsf{B}_{\mathsf{R}}(k)$ (respectively, $\mathsf{C}_{\mathsf{R}}(k)$) as the set of base terms (respectively, canonical terms) of height at most k, then it is clear that $\Delta_{\mathsf{R}}(k)$ is $\underline{\mathsf{B}_{\mathsf{R}}(k)}$ and $\mathsf{R}\langle\Delta_{\mathsf{R}}(k)\rangle$ is $\underline{\mathsf{C}_{\mathsf{R}}(k)}$. Hence, $\Delta_{\mathsf{R}} = \underline{\mathsf{B}_{\mathsf{R}}}$ and $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle = \underline{\mathsf{C}_{\mathsf{R}}} = (\Lambda_{\mathsf{R}}/\triangleq)$. Thus one important contribution of the present paper is to shed new light on the structure of $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ by deliberately introducing α-equivalence and permutative equality separately from the equality of linear combinations (that is, algebraic equality). Also, this gives prominence to the fact that the reduction of the algebraic λ-calculus is defined up to $\triangleq$ (see next section).

So, from now on, we formally identify Δ_{R} with $\underline{\mathsf{B}_{\mathsf{R}}}$ and $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ with $\underline{\mathsf{C}_{\mathsf{R}}}$ by replacing Definition 2.21 with the following one.

Definition 2.22. We define *simple terms* as the $\triangleq$-classes of base terms. We write Δ_{R} for the set of simple terms and $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ for the set of algebraic terms, which we may just call *terms* for short.

When we write a simple term (respectively, a term) as $\underline{s}$, $\underline{t}$, $\underline{u}$, $\underline{v}$ or $\underline{w}$ (respectively, $\underline{S}$, $\underline{T}$, $\underline{U}$, $\underline{V}$ or $\underline{W}$), it is implicit that s, t, u, v, or w is a base term (respectively, S, T, U, V, or W is a canonical term). When we wish to make no such assumption, we write $\underline{L}$, $\underline{M}$ or $\underline{N}$ or use greek letters σ, τ, ρ. We will often use the notations $\lambda x\,\sigma$, $(\sigma)\,\tau$, $a\sigma$, $\sigma+\tau$ with the obvious sense: these are well defined by the contextuality of $\triangleq$.

Definition 2.23. For all $\underline{S} \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ and $\underline{s} \in \Delta_{\mathsf{R}}$, we define the coefficient of $\underline{s}$ in $\underline{S}$ by $\underline{S}_{(\underline{s})} = S_{(s)}$. We then define the support of $\underline{S}$ as the set of all simple terms with a non-zero

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

coefficient in $\underline{S}$:

$$\mathsf{Supp}(\underline{S}) = \left\{\underline{s} \in \Delta_{\mathsf{R}};\ \underline{S}_{(\underline{s})} \neq 0\right\}.$$

If $\mathscr{S}$ is a set of simple terms, we write $\mathsf{R}\langle\mathscr{S}\rangle$ for the set of linear combinations of elements of $\mathscr{S}$, that is,

$$\mathsf{R}\langle\mathscr{S}\rangle = \left\{\underline{\sum_{i=1}^{n} a_i s_i};\ \forall i \in \{1,\ldots,n\},\ \underline{s_i} \in \mathscr{S}, a_i \in \mathsf{R}\right\}$$

or, equivalently, $\mathsf{R}\langle\mathscr{S}\rangle = \{\sigma \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle\ ;\ \mathsf{Supp}(\sigma) \subseteq \mathscr{S}\}$.

3. Reductions

In this section we define reduction using (3) and (4) as key reduction rules, thereby capturing the definition of reduction in Ehrhard and Regnier (2003), minus differentiation, in the setting of the algebraic λ-calculus.

3.1. *Reduction and linear combinations of terms*

We call any subset of $\Delta_{\mathsf{R}} \times \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ a *relation from simple terms to terms*, and any subset of $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle \times \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ a *relation from terms to terms*. Given a relation r from simple terms to terms, we define two new relations $\bar{\mathrm{r}}$ and $\tilde{\mathrm{r}}$ from terms to terms by:

— $\sigma\ \bar{\mathrm{r}}\ \sigma'$ if $\sigma = \underline{\sum_{i=1}^{n} a_i s_i}$ and $\sigma' = \underline{\sum_{i=1}^{n} a_i S'_i}$ where, for all $i \in \{1,\ldots,n\}$, $\underline{s_i}\ \mathrm{r}\ \underline{S'_i}$;
— $\sigma\ \tilde{\mathrm{r}}\ \sigma'$ if $\sigma = \underline{as + T}$ and $\sigma' = \underline{aS' + T}$ where $a \neq 0$ and $\underline{s}\ \mathrm{r}\ \underline{S'}$.

Clearly, $\tilde{\mathrm{r}} \subseteq \bar{\mathrm{r}}$. An important feature of the above definitions is that we do not require $\sum_{i=1}^{n} a_i s_i$ or $as + T$ to be canonical terms: $\tilde{\mathrm{r}}$ matches equation (4), while $\bar{\mathrm{r}}$ matches (5). We will use these constructions in the definitions of one-step β-reduction $\to$ and parallel reduction $\rightrightarrows$. We will introduce these as relations from simple terms to terms so that the actual reduction relations on terms are obtained as $\tilde{\to}$ and $\overline{\rightrightarrows}$, respectively.

Note that we cannot define reduction by induction on terms, since if there are $a, b \in \mathsf{R}^{\bullet}$ such that $a + b = 0$, then $\underline{\mathbf{0}} = a\sigma + b\sigma$ for all $\sigma \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$, and thus by rule (4), $\underline{\mathbf{0}}$ may reduce. Instead, following Ehrhard and Regnier (2003), we define simple term reduction $\to$ by induction on the depth of the fired redex.

Definition 3.1. We define an increasing sequence of relations from simple terms to terms by the following statements. Let $\to_0$ be the empty relation $\varnothing \subseteq \Delta_{\mathsf{R}} \times \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$, and assume that $\to_k$ is defined. Then we set $\sigma \to_{k+1} \sigma'$ when one of the following holds:

— $\sigma = \underline{\lambda x\, s}$ and $\sigma = \underline{\lambda x\, S'}$ with $\underline{s} \to_k \underline{S'}$;
— $\sigma = \underline{(s)\,T}$ and $\sigma' = \underline{(S')\,T}$ with $\underline{s} \to_k \underline{S'}$, or $\sigma' = \underline{(s)\,T'}$ with $\underline{T}\ \widetilde{\to_k}\ \underline{T'}$;
— $\sigma = \underline{(\lambda x\, s)\,T}$ and $\sigma' = \underline{s\,[T/x]}$.

Let $\to\ = \bigcup_{k\in\mathbf{N}} \to_k$. We call the relation $\tilde{\to}$ *one-step reduction* or simply *reduction* for short.

Lemma 3.2. $\tilde{\to} = \bigcup_{k\in\mathbf{N}} \widetilde{\to_k}$.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

Proof. The result is a consequence of the following more general properties of $\widetilde{\cdot}$. If (r_n) is an increasing sequence of relations from simple terms to terms, then $(\widetilde{\mathsf{r}_n})$ is also increasing (monotony) and $\widetilde{\bigcup_n \mathsf{r}_n} = \bigcup_n \widetilde{\mathsf{r}_n}$ (ω-continuity). □

Lemma 3.3. If $\sigma \in \Delta_{\mathsf{R}}$ and $\sigma' \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$, then $\sigma \to \sigma'$ if and only if one of the following holds:

(i) $\sigma = \lambda x\, \tau$ and $\sigma = \lambda x\, \tau'$ with $\tau \to \tau'$;
(ii) $\sigma = (\tau)\, \rho$ and $\sigma' = (\tau')\, \rho$ with $\tau \to \tau'$, or $\sigma' = (\tau)\, \rho'$ with $\rho \mathrel{\widetilde{\to}} \rho'$;
(iii) $\sigma = (\lambda x\, \tau)\, \rho$ and $\sigma' = \tau\,[\rho/x]$;

where $\tau \in \Delta_{\mathsf{R}}$ in each case.

Proof. If (i) or the first case of (ii) holds, it holds at some depth k, hence $\sigma \to_{k+1} \sigma'$. If the second case of (ii) holds, then, by Lemma 3.2, we get $\rho \mathrel{\widetilde{\to_k}} \rho'$ for some k, hence $\sigma \to_{k+1} \sigma'$. If (iii) holds, then $\sigma \to_1 \sigma'$. Conversely, if $\sigma \to \sigma'$, then there is k such that $\sigma \to_k \sigma'$ and one of (i) (ii) or (iii) holds by the definition of $\to_k$ (and Lemma 3.2 in the second case of (ii)). □

Let $\widetilde{\to}^*$ be the reflexive and transitive closure of $\widetilde{\to}$.

Lemma 3.4. Let $\sigma, \sigma' \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ with $\sigma \mathrel{\widetilde{\to}} \sigma'$. Then for all $\tau \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ and all $a \in \mathsf{R}$ we have:

$$\begin{aligned}
\lambda x\, \sigma &\mathrel{\widetilde{\to}} \lambda x\, \sigma' \\
(\sigma)\, \tau &\mathrel{\widetilde{\to}} (\sigma')\, \tau \\
(\tau)\, \sigma &\mathrel{\widetilde{\to}^*} (\tau)\, \sigma' \\
\sigma + \tau &\mathrel{\widetilde{\to}} \sigma' + \tau \\
\sigma &\mathrel{\widetilde{\to}^*} a\sigma'.
\end{aligned}$$

Proof. We write $\sigma = \underline{S} = \underline{bu + V}$ and $\sigma' = \underline{S'} = \underline{bU' + V}$ with $b \neq 0$ and $\underline{u} \to \underline{U'}$, and write $\tau = \underline{T} = \underline{\sum_{i=1}^n a_i t_i}$. Then, by Lemma 3.3, $\underline{\lambda x\, u} \to \underline{\lambda x\, U'}$ and $\underline{(u)\, T} \to \underline{(U')\, T}$. So

$$\lambda x\, \sigma = \underline{b\lambda x\, u + \lambda x\, V} \mathrel{\widetilde{\to}} \underline{b\lambda x\, U' + \lambda x\, V} = \lambda x\, \sigma'$$

and

$$(\sigma)\, \tau = \underline{b(u)\, T + (V)\, T} \mathrel{\widetilde{\to}} \underline{b(U')\, T + (V)\, T} = (\sigma')\, \tau.$$

Also, for each i, we have $\underline{(t_i)\, S} \to \underline{(t_i)\, S'}$. Then, in n $\widetilde{\to}$-steps, $(\tau)\, \sigma = \underline{\sum_{i=1}^n a_i(t_i)\, S}$ reduces to $(\tau)\, \sigma' = \underline{\sum_{i=1}^n a_i(t_i)\, S'}$. For sum, we have $\sigma + \tau = \underline{bu + V + T} \mathrel{\widetilde{\to}} \underline{bU' + V + T} = \sigma' + \tau$. If $ab = 0$, we have $\underline{abu} = \underline{abU'} = \mathbf{0}$, and thus $a\sigma = a\sigma'$, otherwise we have $a\sigma = \underline{abu + aV} \mathrel{\widetilde{\to}} \underline{abU' + aV} = a\sigma'$. □

Lemma 3.5. The relation $\widetilde{\to}^*$ is contextual.

Proof. This is a straightforward consequence of Lemma 3.4 using reflexivity and transitivity. □

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

3.2. *Confluence*

We prove the confluence of $\tilde{\to}$ by the usual Tait–Martin-Löf technique of introducing a parallel extension of reduction (in which redexes can be fired simultaneously) and proving that it enjoys the diamond property (that is, strong confluence).

3.2.1. *Parallel reduction*

Definition 3.6. We define an increasing sequence of relations from simple terms to terms by the following statements. Let $\rightrightarrows_0$ be the identity relation on Δ_{R} extended as a relation from simple terms to terms, and assume that $\rightrightarrows_k$ is defined. Then we set $\sigma \rightrightarrows_{k+1} \sigma'$ when one of the following holds:

— $\sigma = \underline{\lambda x\, s}$ and $\sigma' = \underline{\lambda x\, S'}$ with $\underline{s} \rightrightarrows_k \underline{S'}$;
— $\sigma = \underline{(s)\, T}$ and $\sigma' = \underline{(S')\, T'}$ with $\underline{s} \rightrightarrows_k \underline{S'}$ and $\underline{T}\ \overline{\rightrightarrows_k}\ \underline{T'}$;
— $\sigma = \underline{(\lambda x\, s)\, T}$ and $\sigma' = \underline{S'\,[T'/x]}$ with $\underline{s} \rightrightarrows_k \underline{S'}$ and $\underline{T}\ \overline{\rightrightarrows_k}\ \underline{T'}$.

Let $\rightrightarrows = \bigcup_{k\in\mathbf{N}} \rightrightarrows_k$. We call the relation $\overline{\rightrightarrows}$ *parallel reduction.*

Lemma 3.7. $\overline{\rightrightarrows} = \bigcup_{k\in\mathbf{N}} \overline{\rightrightarrows_k}$.

Proof. The proof is similar to the proof of Lemma 3.2 using the fact that $\overline{\,\cdot\,}$ is monotone and ω-continuous. □

Lemma 3.8. If $\sigma \in \Delta_{\mathsf{R}}$ and $\sigma' \in \mathsf{R}\,\langle\Delta_{\mathsf{R}}\rangle$, then $\sigma \rightrightarrows \sigma'$ if and only if one of the following holds:

(i) $\sigma = \lambda x\, \tau$ and $\sigma = \lambda x\, \tau'$ with $\tau \rightrightarrows \tau'$;
(ii) $\sigma = (\tau)\,\rho$ and $\sigma' = (\tau')\,\rho'$ with $\tau \rightrightarrows \tau'$ and $\rho\ \overline{\rightrightarrows}\ \rho'$;
(iii) $\sigma = (\lambda x\, \tau)\,\rho$ and $\sigma' = \tau'\,[\rho'/x]$ with $\tau \rightrightarrows \tau'$ and $\rho\ \overline{\rightrightarrows}\ \rho'$;

where $\tau \in \Delta_{\mathsf{R}}$ in each case.

Proof. As in Lemma 3.3, this is just a rephrasing of the definition of $\rightrightarrows$, where we use Lemma 3.7 when $\overline{\rightrightarrows}$ is involved. □

Lemma 3.9. The relation $\overline{\rightrightarrows}$ is contextual.

Proof. The proof is very similar to that of Lemma 3.4, using Lemma 3.8 and the definition of $\overline{\rightrightarrows}$. □

Lemma 3.10. $(\lambda x\, \sigma)\,\tau\ \overline{\rightrightarrows}\ \sigma'\,[\tau'/x]$ when $\sigma\ \overline{\rightrightarrows}\ \sigma'$ and $\tau\ \overline{\rightrightarrows}\ \tau'$.

Proof. The statement is a straightforward consequence of Lemmas 3.8 and 3.9. □

Lemma 3.11. The following strict inclusions hold:

$$\tilde{\to} \subset \overline{\rightrightarrows} \subset \tilde{\to}^*.$$

Proof. The fact that $\tilde{\to} \subseteq \overline{\rightrightarrows}$ is straightforward from the definitions. The fact that $\rightrightarrows_k \subseteq \tilde{\to}^*$ and $\overline{\rightrightarrows_k} \subseteq \tilde{\to}^*$ is easily proved by induction on k, so $\overline{\rightrightarrows} \subseteq \tilde{\to}^*$. The inclusions are strict since if we write $I = \lambda x\, x$, we have $\underline{(I)\,(I)\,I}\ \overline{\rightrightarrows}\ \underline{I}$ but $\underline{(I)\,(I)\,I}\ \not\tilde{\to}\ \underline{I}$, and $\underline{((I)\,I)\,I}\ \tilde{\to}^*\ \underline{I}$ but $\underline{((I)\,I)\,I}\ \not\overline{\rightrightarrows}\ \underline{I}$. □

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

3.2.2. *Reductions and substitution* The main property of parallel reduction is given by the following lemma, which fails for one-step reduction.

Lemma 3.12. Let x be a variable and $\sigma, \tau, \sigma', \tau'$ be terms. If $\sigma \overline{\rightrightarrows} \sigma'$ and $\tau \overline{\rightrightarrows} \tau'$, then

$$\sigma\,[\tau/x] \overline{\rightrightarrows} \sigma'\,[\tau'/x].$$

Proof. We prove by induction on k that if $\sigma \overline{\rightrightarrows}_k \sigma'$ and $\tau \overline{\rightrightarrows} \tau'$, then $\sigma\,[\tau/x] \overline{\rightrightarrows} \sigma'\,[\tau'/x]$. If $k = 0$, then $\sigma' = \sigma$. So by Lemma 3.9 and Proposition 2.6, we have $\sigma\,[\tau/x] \overline{\rightrightarrows} \sigma\,[\tau'/x] = \sigma'\,[\tau'/x]$. Now we suppose the result holds for some k and extend it to $k+1$ by inspecting the possible cases for reduction $\sigma \overline{\rightrightarrows}_{k+1} \sigma'$. We first address the case in which σ is simple and $\sigma \rightrightarrows_{k+1} \sigma'$. Then one of the following statements applies (we write $\tau = \underline{T}$ and $\tau' = \underline{T'}$):

— $\sigma = \underline{\lambda y\, u}$ with $y \neq x$ and y not free in T, and $\sigma' = \underline{\lambda y\, U'}$ with $\underline{u} \rightrightarrows_k \underline{U'}$.
Hence, by the induction hypothesis, $\underline{u\,[T/x]} \overline{\rightrightarrows} \underline{U'\,[T'/x]}$ and we get

$$\sigma\,[\tau/x] = \underline{\lambda y\, \left(u\,[T/x]\right)} \overline{\rightrightarrows} \underline{\lambda y\, \left(U'\,[T'/x]\right)} = \sigma'\,[\tau'/x]$$

by Lemma 3.9.

— $\sigma = \underline{(u)\,V}$ and $\sigma' = \underline{(U')\,V'}$ with $\underline{u} \rightrightarrows_k \underline{U'}$ and $\underline{V} \overline{\rightrightarrows}_k \underline{V'}$.
Hence, by the induction hypothesis, $\underline{u\,[T/x]} \overline{\rightrightarrows} \underline{U'\,[T'/x]}$ and $\underline{V\,[T/x]} \overline{\rightrightarrows} \underline{V'\,[T'/x]}$, and we get

$$\sigma\,[\tau/x] = \underline{(u\,[T/x])\,V\,[T/x]} \overline{\rightrightarrows} \underline{(U'\,[T'/x])\,V'\,[T'/x]} = \sigma'\,[\tau'/x]$$

by Lemma 3.9.

— $\sigma = \underline{(\lambda y\, u)\,V}$ and $\sigma' = \underline{U'\,[V'/y]}$ with $\underline{u} \rightrightarrows_k \underline{U'}$, $\underline{V} \overline{\rightrightarrows}_k \underline{V'}$, $x \neq y$ and y not free in T.
Hence, by the induction hypothesis, $\underline{u\,[T/x]} \overline{\rightrightarrows} \underline{U'\,[T'/x]}$ and $\underline{V\,[T/x]} \overline{\rightrightarrows} \underline{V'\,[T'/x]}$, and we get

$$\sigma\,[\tau/x] = \underline{(\lambda y\, u\,[T/x])\,V\,[T/x]} \overline{\rightrightarrows} \underline{\left(U'\,[T'/x]\right)\,[V'\,[T'/x]/y]} = \sigma'\,[\tau'/x]$$

by Lemma 3.10.

Now assume $\sigma \overline{\rightrightarrows}_{k+1} \sigma'$. By definition, this amounts to $\sigma = \underline{\sum_{i=1}^n a_i s_i}$ and $\sigma' = \underline{\sum_{i=1}^n a_i S'_i}$, with $\underline{s_i} \rightrightarrows_{k+1} \underline{S'_i}$ for all i. We have just shown that we then have $\underline{s_i\,[T/x]} \overline{\rightrightarrows} \underline{S'_i\,[T'/x]}$ for all i. Lemma 3.9 then gives the required result. □

From Lemmas 3.11 and 3.12, we can derive a very similar result for $\tilde{\rightarrow}^*$.

Corollary 3.13. Let x be a variable and $\sigma, \tau, \sigma', \tau'$ be terms. If $\sigma \tilde{\rightarrow}^* \sigma'$ and $\tau \tilde{\rightarrow}^* \tau'$, then

$$\sigma\,[\tau/x] \tilde{\rightarrow}^* \sigma'\,[\tau'/x].$$

3.2.3. *Church–Rosser* We now conclude the proof of confluence by showing that the $\overline{\rightrightarrows}$-reducts of a fixed term σ all $\overline{\rightrightarrows}$-reduce to one of them (which is obtained by firing all redexes of σ simultaneously).

Definition 3.14. We define inductively on term σ its full parallel reduct $\sigma{\downarrow}$ by:

$$
\begin{aligned}
\underline{x}{\downarrow} &= \underline{x} \\
\underline{\lambda x\, s}{\downarrow} &= \lambda x\ \underline{s}{\downarrow} \\
\underline{(\lambda x\, s)\, T}{\downarrow} &= (\underline{s}{\downarrow})\,[\underline{T}{\downarrow}/x] \\
\underline{(s)\, T}{\downarrow} &= (\underline{s}{\downarrow})\ \underline{T}{\downarrow} \text{ if } s \text{ is a variable or an application} \\
\left.\underline{\sum_{i=1}^{n} a_i s_i}\right|_{\downarrow} &= \sum_{i=1}^{n} a_i\ \underline{s_i}{\downarrow}\,.
\end{aligned}
$$

Lemma 3.15. If σ and σ' are such that $\sigma \mathrel{\overline{\rightrightarrows}} \sigma'$, then $\sigma' \mathrel{\overline{\rightrightarrows}} \sigma{\downarrow}$.

Proof. One simply proves by induction on k that if $\sigma \rightrightarrows_k \sigma'$ or $\sigma \mathrel{\overline{\rightrightarrows}_k} \sigma'$, then $\sigma' \mathrel{\overline{\rightrightarrows}} \sigma{\downarrow}$, using Lemma 3.9 in general, and Lemma 3.10 in the case of a redex. □

Theorem 3.16. Relation $\overline{\rightrightarrows}$ is strongly confluent. Hence, relation $\widetilde{\rightarrow}$ enjoys the Church–Rosser property.

Proof. Strong confluence of $\overline{\rightrightarrows}$ is a straightforward corollary of Lemma 3.15. It implies confluence of $\widetilde{\rightarrow}$ by Lemma 3.11. □

3.2.4. *Trivia* There is a case in which confluence is much easier to establish: if 1 admits an opposite $-1 \in \mathsf{R}$. In this case, assume $\sigma \mathrel{\widetilde{\rightarrow}^*} \sigma'$. Since $\widetilde{\rightarrow}^*$ is contextual, $\sigma' = \sigma' + (-1)\sigma + \sigma \mathrel{\widetilde{\rightarrow}^*} \sigma' + (-1)\sigma' + \sigma = \sigma$. Hence, $\widetilde{\rightarrow}^*$ is symmetric, which obviously implies Church–Rosser. But this has little meaning since in the next section we will show that reduction becomes trivial when $-1 \in \mathsf{R}$.

3.3. *Conservativity*

Every ordinary λ-term is also a raw term of the algebraic λ-calculus, whose $\triangleq$-class is simple. Let Λ denote the set of all λ-terms and $\rightarrow_\Lambda$ denote the usual β-reduction of the λ-calculus. It is then clear, for all $s, s' \in \Lambda$, that $s \rightarrow_\Lambda s'$ implies $\underline{s} \rightarrow \underline{s'}$. We use $\leftrightarrow$ to denote the reflexive, symmetric and transitive closure of $\widetilde{\rightarrow}$, and $\leftrightarrow_\Lambda$ to denote the usual β-equivalence of the λ-calculus.

Lemma 3.17. The algebraic λ-calculus preserves the equalities of the λ-calculus, that is, for all λ-terms s and t, we have $s \leftrightarrow_\Lambda t$ implies $\underline{s} \leftrightarrow \underline{t}$.

Proof. The statement is a straightforward consequence of the confluence of $\rightarrow_\Lambda$ and the fact that $\rightarrow_\Lambda \subset \widetilde{\rightarrow}$. □

One may wonder if the reverse also holds, that is, if equivalence classes of λ-terms in the algebraic λ-calculus are the same as in the ordinary λ-calculus. If R is $\mathbf{N}$, then $\widetilde{\rightarrow}$-reductions from λ-terms are exactly $\rightarrow_\Lambda$-reductions (restricted to λ-terms, $\triangleq$ then only amounts to α-conversion), and the result holds by the same argument as in Lemma 3.17. In the general case, however, a λ-term does not necessarily reduce to another λ-term, hence the proof is not as easy.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

3.3.1. *The positive case* Recall that a rig R is said to be positive if, for all $a, b \in \mathsf{R}$, we have $a + b = 0$ implies $a = b = 0$. For this case we will prove that for all $s, s' \in \Lambda$, we have $\underline{s} \leftrightarrow \underline{s'}$ implies $s \leftrightarrow_\Lambda s'$ (Theorem 3.24).

Definition 3.18. We define $\Lambda : \mathsf{R}\langle\Delta_\mathsf{R}\rangle \longrightarrow \mathscr{P}(\Lambda)$ by induction on terms:

$$\begin{aligned}
\Lambda(\underline{x}) &= \{x\}\\
\Lambda(\underline{\lambda x\, s}) &= \{\lambda x\, u;\ u \in \Lambda(\underline{s})\}\\
\Lambda\left(\underline{(s)\, T}\right) &= \{(u)\, v;\ u \in \Lambda(\underline{s}) \text{ and } v \in \Lambda(\underline{T})\}\\
\Lambda\left(\underline{\sum_{i=1}^n a_i s_i}\right) &= \bigcup_{i=1}^n \Lambda\left(\underline{s_i}\right).
\end{aligned}$$

The crucial point in that definition is that the sum $\sum_{i=1}^n a_i s_i$ being canonical entails that, for all i, $a_i \neq 0$.

Proposition 3.19. If $s \in \Lambda$, then $\Lambda(\underline{s}) = \{s\}$.

Lemma 3.20. If R is positive and terms $\sigma \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ and $\sigma' \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ are such that $\sigma \mathrel{\tilde{\to}} \sigma'$, then for all $s' \in \Lambda\left(\sigma'\right)$, either $s' \in \Lambda(\sigma)$ or there exists $s \in \Lambda(\sigma)$ such that $s \to_\Lambda s'$.

Proof. The proof is by induction on the depth of the reduction $\sigma \mathrel{\tilde{\to}} \sigma'$, that is, the least k such that $\sigma \mathrel{\tilde{\to}_k} \sigma'$. All induction steps are straightforward, except for the extension from $\to_k$ to $\tilde{\to}_k$. For this we assume $\sigma = \underline{at + U}$ and $\sigma = \underline{aT' + U}$ with $a \neq 0$ and $\underline{t} \to_k \underline{T'}$. By definition, $\Lambda\left(\sigma'\right) = \Lambda\left(\underline{aT' + U}\right) \subseteq \Lambda\left(\underline{T'}\right) \cup \Lambda(\underline{U})$. Since R is positive, the coefficient of t in $\mathsf{can}(at + U)$ is non-zero, so $\Lambda(\sigma) = \Lambda\left(\underline{at + U}\right) = \Lambda(\underline{t}) \cup \Lambda(\underline{U})$. Now assume $v' \in \Lambda\left(\sigma'\right)$. Either $v' \in \Lambda(\underline{U}) \subseteq \Lambda(\sigma)$ or $v' \in \Lambda\left(\underline{T'}\right)$, in which case, by the induction hypothesis, either $v' \in \Lambda(\underline{t}) \subseteq \Lambda(\sigma)$ or there exists $v \in \Lambda(\underline{t}) \subseteq \Lambda(\sigma)$ such that $v \to_\Lambda v'$. □

Corollary 3.21. If R is positive and $s \in \Lambda$ and $\sigma \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ are such that $\underline{s} \mathrel{\tilde{\to}^*} \sigma$, then for all $t \in \Lambda(\sigma)$, we have $s \to_\Lambda^* t$.

Lemma 3.22. If σ and $\sigma' \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ are such that $\sigma \mathrel{\overline{\rightrightarrows}} \sigma'$, then $\sigma{\downarrow} \mathrel{\overline{\rightrightarrows}} \sigma'{\downarrow}$.

Proof. The proof is easy and very close to that of Lemma 3.15. □

We define iterated full reduction by $\sigma{\downarrow}^0 = \sigma$ and $\sigma{\downarrow}^{n+1} = (\sigma{\downarrow}^n){\downarrow}$.

Lemma 3.23. If $\sigma \mathrel{\overline{\rightrightarrows}^n} \tau$, then $\tau \mathrel{\tilde{\to}^*} \sigma{\downarrow}^n$.

Proof. The proof is by induction on n.

If $n = 0$, we have $\sigma = \tau = \sigma{\downarrow}^0$, which is reflexivity of $\tilde{\to}^*$.

Now assume the result holds at rank n. If $\sigma \mathrel{\overline{\rightrightarrows}^n} \tau \mathrel{\overline{\rightrightarrows}} \tau'$, then, by the induction hypothesis, $\tau \mathrel{\tilde{\to}^*} \sigma{\downarrow}^n$. Since $\tilde{\to}^*$ is also the transitive closure of $\overline{\rightrightarrows}$, Lemma 3.22 entails $\tau{\downarrow} \mathrel{\tilde{\to}^*} \sigma{\downarrow}^{n+1}$. By Lemma 3.15, we have $\tau' \mathrel{\overline{\rightrightarrows}} \tau{\downarrow}$, hence $\tau' \mathrel{\tilde{\to}^*} \sigma{\downarrow}^{n+1}$. □

Theorem 3.24. If R is positive and $s, t \in \Lambda$ are such that $s \leftrightarrow t$, then $s \leftrightarrow_\Lambda t$.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

$$\frac{}{\Gamma, x : A \vdash x : A} \text{(Ax)}$$

$$\frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x\, M : A \to B} \text{(Abs)} \qquad \frac{\Gamma \vdash M : A \to B \qquad \Gamma \vdash N : A}{\Gamma \vdash (M)\, N : B} \text{(App)}$$

$$\frac{}{\Gamma \vdash \mathbf{0} : A} \text{(Zero)} \qquad \frac{\Gamma \vdash M : A}{\Gamma \vdash aM : A} \text{(Scal)} \qquad \frac{\Gamma \vdash M : A \qquad \Gamma \vdash N : A}{\Gamma \vdash M + N : A} \text{(Add)}$$

Figure 1. Typing rules for the algebraic λ-calculus.

Proof. Assume $s, t \in \Lambda$ and $\underline{s} \leftrightarrow \underline{t}$. By the Church–Rosser property of $\widetilde{\to}$ (Theorem 3.16), there exists $\sigma \in \mathsf{R}\langle \Delta_{\mathsf{R}} \rangle$ such that $\underline{s} \widetilde{\to}^* \sigma$ and $\underline{t} \widetilde{\to}^* \sigma$. By Lemma 3.23, there exists some $n \in \mathbf{N}$ such that $\sigma \widetilde{\to}^* \tau = \underline{s}{\downarrow}^n$. Note that for all $\underline{w} \in \underline{\Lambda}$, we have $\underline{w}{\downarrow} \in \underline{\Lambda}$. So $\tau \in \underline{\Lambda}$ and we write $\tau = \underline{v}$ with $v \in \Lambda$. We have $\underline{s} \widetilde{\to}^* \underline{v}$ and $\underline{t} \widetilde{\to}^* \underline{v}$: by positivity of R and Corollary 3.21, we obtain that, for all $v' \in \Lambda(\underline{v})$, there are $s' \in \Lambda(\underline{s})$ and $t' \in \Lambda(\underline{t})$ such that $s' \to_\Lambda^* v'$ and $t' \to_\Lambda^* v'$. By Proposition 3.19, $\Lambda(\underline{s}) = \{s\}$, $\Lambda(\underline{t}) = \{t\}$ and $\Lambda(\underline{v}) = \{v\}$, which concludes the proof. □

3.3.2. *Collapse* If R is not positive, we show that reductional equality collapses since $\leftrightarrow$ identifies terms that bear absolutely no relationship to each other.

Lemma 3.25. Assume, there are $a, b \in \mathsf{R}^\bullet$ such that $a + b = 0$. Then for any term σ, $\underline{\mathbf{0}} \widetilde{\to}^* a\sigma \widetilde{\to}^* \underline{\mathbf{0}}$.

Proof. Take Θ to be a fixed point combinator of the λ-calculus such that we have $(\Theta)\, s \to_\Lambda^* (s)\,(\Theta)\, s$ for all λ-terms s. We write ∞_σ for $(\Theta)\, \lambda x\, (\sigma + \underline{x})$, so we have $\infty_\sigma \widetilde{\to}^* \sigma + \infty_\sigma$. We then get

$$\underline{\mathbf{0}} = a\infty_\sigma + b\infty_\sigma \widetilde{\to}^* a\sigma + a\infty_\sigma + b\infty_\sigma = a\sigma$$

and

$$a\sigma = a\sigma + a\infty_\sigma + b\infty_\sigma \widetilde{\to}^* a\sigma + a\infty_\sigma + b\sigma + b\infty_\sigma = \underline{\mathbf{0}}.$$ □

Corollary 3.26. If R is such that 1 has an opposite, that is, $-1 \in \mathsf{R}$ with $1 + (-1) = 0$, then for all terms σ and τ, we have $\sigma \widetilde{\to}^* \tau$.

4. Simple type system

Raw terms may be given implicative propositional types in a natural way. Assume we have a denumerable set of basic types $\phi, \psi, \ldots$. We build types from basic types using intuitionistic arrow: if A and B are types, then so is $A \to B$. Typing rules are given in Figure 1. Notice that scalar coefficients have no influence on typing. In particular, we make no assumption on the actual structure of R.

Proposition 4.1. Typing in the algebraic λ-calculus enjoys the following properties:

(i) If $\Gamma \vdash M : A$, then free variables of M are declared in Γ.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

(ii) If $\Gamma \vdash M : A$, then, for all Γ' whose domain is disjoint from that of Γ, we have $\Gamma, \Gamma' \vdash M : A$.
(iii) If $M \equiv M'$, then $\Gamma \vdash M : A$ if and only if $\Gamma \vdash M' : A$.
(iv) For all canonical terms S, we have $\Gamma \vdash S : A$ if and only if for all $u \in \Lambda(\underline{S})$, we have $\Gamma \vdash u : A$.
(v) For all raw terms M, if $\Gamma \vdash M : A$, then $\Gamma \vdash \mathsf{can}(M) : A$.

The converse of the last part of the proposition does not hold. For instance, for all raw terms M, $\mathsf{can}(0M) = \mathbf{0}$ can be given any type in any context whereas $0M$ satisfies the same typing judgements as M. Hence, such a straightforward notion of typing is not compatible with algebraic equality $\triangleq$.

Definition 4.2. The term σ is *weakly typable* of type A in context Γ if $\Gamma \vdash \mathsf{can}(\sigma) : A$ is derivable. We write $\Gamma \vdash_{\mathsf{R}} \sigma : A$ for $\Gamma \vdash \mathsf{can}(\sigma) : A$.

Proposition 4.3. For all $\sigma \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$, we have $\Gamma \vdash_{\mathsf{R}} \sigma : A$ if and only if $\Gamma \vdash_{\mathsf{R}} \underline{s} : A$ for all $\underline{s} \in \mathsf{Supp}(\sigma)$.

We now show that subject reduction holds for weak typing when R is positive (culminating in Lemma 4.6).

Lemma 4.4. Let $\sigma, \tau \in \Lambda_{\mathsf{R}}$. If $\Gamma, x : A \vdash_{\mathsf{R}} \sigma : B$ and $\Gamma \vdash_{\mathsf{R}} \tau : A$, then $\Gamma \vdash_{\mathsf{R}} \sigma[\tau/x] : B$.

Proof. We prove by induction on the derivation of $\Gamma, x : A \vdash M : B$ that if we also have $\Gamma \vdash N : A$, then $\Gamma \vdash M[N/x] : B$. The result follows by taking $M = \mathsf{can}(\sigma)$ and $N = \mathsf{can}(\tau)$, using Lemma 2.16 (v). □

Lemma 4.5. For all $\sigma, \tau \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ and all $a \in \mathsf{R}$, we have $\mathsf{Supp}(\sigma + \tau) \subseteq \mathsf{Supp}(\sigma) \cup \mathsf{Supp}(\tau)$ and $\mathsf{Supp}(a\sigma) \subseteq \mathsf{Supp}(\sigma)$. If R is positive, we also have $\mathsf{Supp}(\sigma + \tau) = \mathsf{Supp}(\sigma) \cup \mathsf{Supp}(\tau)$.

Proof. For all $\underline{s} \in \Delta_{\mathsf{R}}$, we have $(\sigma + \tau)_{(\underline{s})} = \sigma_{(\underline{s})} + \tau_{(\underline{s})}$ and $(a\sigma)_{(\underline{s})} = a\sigma_{(\underline{s})}$. By the definition of $\mathsf{Supp}(\sigma + \tau)$ and $\mathsf{Supp}(a\sigma)$, we get the above inclusions. If R is positive, $(\sigma + \tau)_{(\underline{s})} \neq 0$ when $\sigma_{(\underline{s})} \neq 0$ or $\tau_{(\underline{s})} \neq 0$, hence $\mathsf{Supp}(\sigma + \tau) = \mathsf{Supp}(\sigma) \cup \mathsf{Supp}(\tau)$. □

Notice that we do not necessarily have $\mathsf{Supp}(a\sigma) = \mathsf{Supp}(\sigma)$ when $a \neq 0$ and R is positive: see Lemma 5.3 for a sufficient condition.

Lemma 4.6. Assume R is positive. If $\sigma \tilde{\to} \sigma'$ and $\Gamma \vdash_{\mathsf{R}} \sigma : A$, then $\Gamma \vdash_{\mathsf{R}} \sigma' : A$.

Proof. We prove by induction on base and canonical terms that if either $\Gamma \vdash s : A$ and $\underline{s} \to \sigma'$, or $\Gamma \vdash S : A$ and $\underline{S} \tilde{\to} \sigma'$, then $\Gamma \vdash_{\mathsf{R}} \sigma' : A$.

For base terms, we check that all possible cases for reduction $\underline{s} \to \sigma'$ preserve weak typing, which is straightforward by the induction hypotheses (using Lemma 4.4 in the case of a redex).

Now assume $\Gamma \vdash S : A$ and write $\underline{S} = \underline{at + U}$ and $\sigma' = \underline{aT' + U}$, with $a \neq 0$ and $\underline{t} \to \underline{T'}$. By Lemma 4.5, $\mathsf{Supp}(\underline{S}) = \{\underline{t}\} \cup \mathsf{Supp}(\underline{U})$ (this is where we use the positivity condition). By Proposition 4.3, $\Gamma \vdash t : A$ and $\Gamma \vdash U : A$. By the induction hypothesis on base term t, we get $\Gamma \vdash T' : A$. By Lemma 4.5 again, $\mathsf{Supp}(\sigma') \subseteq \mathsf{Supp}(\underline{T'}) \cup \mathsf{Supp}(\underline{U})$, and we get $\Gamma \vdash_{\mathsf{R}} \sigma' : A$ by Proposition 4.3. □

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

5. On normalisation properties

Unsurprisingly, if R is not positive, there is no normal term. To see this, assume there are $a, b \in \mathsf{R}^\bullet$ such that $a + b = 0$ and let $\sigma \in \Delta_\mathsf{R}$ and $\sigma' \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ be such that $\sigma \to \sigma'$. Then for all $\tau \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$, we have $\tau = a\sigma + b\sigma + \tau$, so $\tau \mathrel{\tilde{\to}} a\sigma' + b\sigma + \tau$. Hence every term τ reduces.

Moreover, even if R is positive, it may be the case that the only normalisable terms are normal terms. Indeed, assume R is the set $\mathbf{Q}^+$ of non-negative rational numbers (which is a positive rig), and $\sigma \to \sigma'$. Then there is an infinite sequence of reductions from σ:

$$\sigma = \frac{1}{2}\sigma + \frac{1}{2}\sigma \mathrel{\tilde{\to}} \frac{1}{2}\sigma + \frac{1}{2}\sigma' \mathrel{\tilde{\to}} \frac{1}{4}\sigma + \frac{3}{4}\sigma' \mathrel{\tilde{\to}} \cdots \mathrel{\tilde{\to}} \frac{1}{2^n}\sigma + \frac{2^n - 1}{2^n}\sigma' \mathrel{\tilde{\to}} \cdots$$

In order to establish the strong normalisation of typed terms, we will therefore assume that R is *finitely splitting* in the sense that for all $a \in \mathsf{R}$,

$$\{(a_1, \ldots, a_n) \in (\mathsf{R}^\bullet)^n \ ; \ n \in \mathbf{N} \text{ and } a = a_1 + \cdots + a_n\}$$

is finite. We can then define the *width* function

$$\mathsf{w}(a) = \max\{n \in \mathbf{N};\ \exists (a_1, \ldots, a_n) \in (\mathsf{R}^\bullet)^n \text{ such that } a = a_1 + \cdots + a_n\}.$$

The width function relates the additive structure of R to that of $\mathbf{N}$, as shown by the following lemma.

Lemma 5.1. If R is finitely splitting, then it is positive. Moreover, for all $a, b \in \mathsf{R}$, we have $\mathsf{w}(a) = 0$ if and only if $a = 0$ and $\mathsf{w}(a + b) \geqslant \mathsf{w}(a) + \mathsf{w}(b)$.

Proof. Assume R is finitely splitting. Since 0 is neutral for addition in R, the empty sequence is the only element of $\{(a_1, \ldots, a_n) \in (\mathsf{R}^\bullet)^n \ ; \ n \in \mathbf{N} \text{ and } a_1 + \cdots + a_n = 0\}$. Hence $\mathsf{w}(0) = 0$ and R is positive. If $a \neq 0$, we have $\mathsf{w}(a) \geqslant 1$. Hence $\mathsf{w}(a) = 0$ implies $a = 0$. Now let $a, b \in \mathsf{R}$. We can write $a = a_1 + \cdots + a_{\mathsf{w}(a)}$ and $b = b_1 + \cdots + b_{\mathsf{w}(b)}$, where the a_i's and the b_j's are non-zero. Then $a + b = a_1 + \cdots + a_{\mathsf{w}(a)} + b_1 + \cdots + b_{\mathsf{w}(b)}$, and thus $\mathsf{w}(a + b) \geqslant \mathsf{w}(a) + \mathsf{w}(b)$. □

An essential point of this section is to show that the finite splitting condition is efficient in preventing the tricky situations we have just seen in $\mathbf{Q}^+$. We are led to prove that strongly normalising terms are exactly the linear combinations of strongly normalising simple terms.

The finite splitting property is actually not sufficient for this. Take, for instance, $\mathsf{R} = \mathbf{N} \times \mathbf{N}$, with operations defined pointwise:

$$\begin{aligned} (p, q) + (p', q') &= (p + p', q + q') \\ (p, q)(p', q') &= (pp', qq'). \end{aligned}$$

It is easy to check that this defines a finitely splitting rig, with $\mathsf{w}(p, q) = p + q$. Now write $a = (1, 0)$ and $b = (0, 1)$. We have $\mathsf{w}(a) = \mathsf{w}(b) = 1$, $a + b = (1, 1) = 1_\mathsf{R}$ and $ab = (0, 0) = 0_\mathsf{R}$. Then if we write $\delta = \lambda x\,(x)\,x$, we may note that the only $\tilde{\to}$-reduct of term $\underline{a(\delta)\,b\delta}$ is $\underline{\mathbf{0}}$, which is normal, whereas the simple term $\underline{(\delta)\,b\delta}$ has no normal form.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

We will therefore require R to be finitely splitting *and* to satisfy the following *integral domain* property. For all $a, b \in \mathsf{R}$, if $ab = 0$, then either $a = 0$ or $b = 0$. When this is the case, we obtain the following four lemmas.

Lemma 5.2. For all $a, b \in \mathsf{R}$, we have $\mathsf{w}(ab) \geqslant \mathsf{w}(a)\,\mathsf{w}(b)$. In particular, $\mathsf{w}(1) = 1$.

Proof. Write $a = a_1 + \cdots + a_{\mathsf{w}(a)}$ and $b = b_1 + \cdots + b_{\mathsf{w}(b)}$, where the a_i's and the b_j's are non-zero. Then, developing $ab = (a_1 + \cdots + a_{\mathsf{w}(a)})(b_1 + \cdots + b_{\mathsf{w}(b)})$, we obtain $\mathsf{w}(a)\,\mathsf{w}(b)$ summands, which are all non-zero by the integral domain property of R. □

Lemma 5.3. If $\sigma = a\tau + \rho$ with $a \neq 0$, then $\mathsf{Supp}(\sigma) = \mathsf{Supp}(\tau) \cup \mathsf{Supp}(\rho)$.

Proof. By Lemma 4.5, all that remains to be shown is that $\mathsf{Supp}(a\tau) = \mathsf{Supp}(\tau)$, which follows directly from the integral domain property of R. □

Lemma 5.4. For all σ, σ' such that $\sigma \tilde{\to} \sigma'$, we have $a\sigma + \tau \tilde{\to} a\sigma' + \tau$ also holds when $a \neq 0$.

Proof. This is again a direct consequence of the integral domain property of R. □

Lemma 5.5. For all $\sigma \in \Delta_{\mathsf{R}}$ and all $\sigma' \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$, we have $\sigma \tilde{\to} \sigma'$ if and only if $\sigma \to \sigma'$.

Proof. By Lemma 5.3 and the fact that $\mathsf{Supp}(\sigma) = \{\sigma\}$, if we write $\sigma = \underline{as + T}$ with $a \neq 0$, then $\underline{s} = \sigma$ and there is $b \in \mathsf{R}$ such that $\underline{T} = b\sigma$. Necessarily, we have $a + b = 1$, which by Lemma 5.2 implies $a = 1$ and $b = 0$, and the result then follows from the definition of $\tilde{\to}$. □

In Subsection 5.1, we will prove that, under these conditions, $\sigma \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ is strongly normalising if and only if every simple term in $\mathsf{Supp}(\sigma)$ is strongly normalising. We then develop the proof of strong normalisation of simply typed terms, in Subsections 5.2–5.4, following Krivine's version of Tait's reducibility method (Krivine 1990). From this, we derive a weak normalisation result in Subsection 5.5 in which the only assumption is that R is positive.

Examples Obviously, the rig $\mathbf{N}$ is finitely splitting with $\mathsf{w}(n) = n$ for all $n \in \mathbf{N}$, and has no zero divisor. One more interesting instance is the rig of all polynomials over variables $\xi_1, \ldots, \xi_n$ with non-negative integer coefficients, which is denoted by $\mathsf{P}_n = \mathbf{N}[\xi_1, \ldots, \xi_n]$. For all $P \in \mathsf{P}_n$, we have $\mathsf{w}(P) = P(1, \ldots, 1)$. Such a rig of polynomials is involved in the weak normalisation scheme we will develop in Section 5.5. All other examples we know of are given by variants of P_n, such as:

— any rig $\mathsf{R}[\xi_1, \ldots, \xi_n]$, where R is itself an integral finitely splitting rig;
— any similar rig of polynomials, with the restriction that the ξ_i's do not commute, that is, $\xi_i\xi_j \neq \xi_j\xi_i$ when $i \neq j$ (this is a rig that satisfies our conditions, but is not commutative for multiplication);
— any similar rig of polynomials, but where the ξ_i's are supposed to be idempotent, that is, $\xi_i\xi_i = \xi_i$ for all i.

5.1. *Scalars and normalisation*

From now on, we assume R is finitely splitting and integral. Under these conditions, we prove a term is strongly normalising if and only if it is a linear combination of strongly normalising simple terms (Theorem 5.11).

Lemma 5.6. Let $\sigma \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$. There are only finitely many terms σ' such that $\sigma \tilde{\to} \sigma'$.

Proof. The proof is by induction on $\mathrm{h}(\sigma)$. If $\mathrm{h}(\sigma) = 0$, then $\sigma = \underline{\mathbf{0}}$ and the property holds trivially by Lemma 4.5. Assume that the property holds for all σ such that $\mathrm{h}(\sigma) \leqslant k$. Let $\sigma \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ be such that $\mathrm{h}(\sigma) = k+1$. For each term $\sigma' \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ such that $\sigma \tilde{\to} \sigma'$, there are $\underline{t} \in \Delta_\mathsf{R}$, $\underline{T'}, \underline{U} \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ and $a \in \mathsf{R}^\bullet$ such that $\sigma = \underline{at + U}$, $\sigma' = \underline{aT' + U}$ and $\underline{t} \to \underline{T'}$. By Lemma 4.5, $\underline{t} \in \mathsf{Supp}(\sigma)$: there are finitely many such simple terms. Moreover, because of the finite splitting condition on R, for each such $\underline{t}$ there exist finitely many $a \in \mathsf{R}^\bullet$ and $\underline{U} \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ such that $\sigma = \underline{at + U}$. A simple inspection of the definition of $\to$ shows that, by the induction hypothesis applied to subterms of $\underline{t}$ (that is, $\triangleq$-classes of subterms of t, all of height at most k), $\underline{t}$ $\to$-reduces to finitely many terms, which are all the possible choices for $\underline{T'}$. □

König's lemma thus justifies the following definition.

Definition 5.7. If σ is a strongly normalising term, we use $|\sigma|$ to denote the length of the longest sequence of $\tilde{\to}$-reductions from σ to its normal form. We use N_R to denote the set of strongly normalising simple terms and $\mathsf{N}_\mathsf{R}(n) = \{\sigma \in \mathsf{N}_\mathsf{R} \text{ such that } |\sigma| \leqslant n\}$.

Then $\mathsf{R}\langle\mathsf{N}_\mathsf{R}\rangle$ is the set of linear combinations of strongly normalising simple terms:

$$\mathsf{R}\langle\mathsf{N}_\mathsf{R}\rangle = \{\sigma \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle\ ;\ \mathsf{Supp}(\sigma) \subseteq \mathsf{N}_\mathsf{R}\}.$$

In the following, we prove that $\mathsf{R}\langle\mathsf{N}_\mathsf{R}\rangle$ is exactly the set of all strongly normalising terms. We first show the easiest inclusion.

Lemma 5.8. The support of every strongly normalising term is a finite subset of N_R. More precisely, if σ is strongly normalising, then $\mathsf{Supp}(\sigma) \subset \mathsf{N}_\mathsf{R}(|\sigma|)$.

Proof. By Lemma 5.4, from a sequence of reductions from $\tau \in \mathsf{Supp}(\sigma)$, we can derive a sequence of reductions from σ of the same length. □

We now establish the reverse inclusion and show that the terms in $\mathsf{R}\langle\mathsf{N}_\mathsf{R}\rangle$ are strongly normalising. The proof boils down to the following idea. To each $\sigma \in \mathsf{R}\langle\mathsf{N}_\mathsf{R}\rangle$, we associate a finite multiset $\|\sigma\|$ of natural numbers so that if $\sigma \tilde{\to} \sigma'$, then $\|\sigma\| > \|\sigma'\|$, where $>$ denotes the *multiset order* (which is a well-order).

First we fix the notation for multisets. We write $\mathscr{M}_{\mathrm{fin}}(\mathbf{N})$ for the set of finite multisets of natural numbers. If $p_1, \ldots, p_n \in \mathbf{N}$, we write $[p_1, \ldots, p_n] \in \mathscr{M}_{\mathrm{fin}}(\mathbf{N})$ for the multiset containing exactly $p_1, \ldots, p_n$, taking repetitions into account. If $\mu, \nu \in \mathscr{M}_{\mathrm{fin}}(\mathbf{N})$, we use $\mu + \nu$ to denote the multiset union of μ and ν, and if $k \in \mathbf{N}$, we use $k\mu$ to denote the multiset $\sum_{i=1}^{k} \mu$. Now assume $\mu = [p_1, \ldots, p_m]$ and $\nu = [q_1, \ldots, q_n]$, with $p_1 \leqslant \cdots \leqslant p_m$ and $q_1 \leqslant \cdots \leqslant q_n$. We recall that $\mu < \nu$ for the multiset order if and only if one of the following holds:

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

— $m = 0$ and $n > 0$;
— $mn \neq 0$ and $p_m < q_n$;
— $mn \neq 0$, $p_m = q_n$ and $[p_1, \ldots, p_{m-1}] < [q_1, \ldots, q_{n-1}]$.

This strict order is the transitive closure of the relation defined by $\mu \prec \mu'$ if and only if $\mu = v + [p_1, \ldots, p_m]$ and $\mu' = v + [q]$ where $p_i < q$ for all i. The well-foundedness of the multiset order amounts to the fact that there is no infinite descending chain for $\prec$.

Definition 5.9. For all $\tau \in \Delta_{\mathsf{R}}$ and $\sigma \in \mathsf{R}\langle \Delta_{\mathsf{R}} \rangle$, we write $\mathsf{w}_\tau(\sigma)$ for the width of the coefficient of τ in σ: $\mathsf{w}_\tau(\sigma) = \mathsf{w}\left(\sigma_{(\tau)}\right)$. If, moreover, $\sigma \in \mathsf{R}\langle \mathsf{N}_{\mathsf{R}} \rangle$, we write

$$\|\sigma\| = \sum_{\tau \in \mathsf{Supp}(\sigma)} \mathsf{w}_\tau(\sigma) \left[|\tau| \right].$$

For instance, if σ is a strongly normalising simple term, $\|\sigma\| = \mathsf{w}(1)\left[|\sigma|\right] = \left[|\sigma|\right]$.

Lemma 5.10. Let $\sigma \in \mathsf{R}\langle \mathsf{N}_{\mathsf{R}} \rangle$ and let σ' be such that $\sigma \to \sigma'$. Then $\sigma' \in \mathsf{R}\langle \mathsf{N}_{\mathsf{R}} \rangle$ and $\|\sigma'\| < \|\sigma\|$.

Proof. We write $\sigma = \underline{as + T}$ and $\sigma' = \underline{aS' + T}$ with $\underline{s} \to \underline{S'}$. Since $\sigma \in \mathsf{R}\langle \mathsf{N}_{\mathsf{R}} \rangle$, Lemma 5.3 entails $\underline{s} \in \mathsf{N}_{\mathsf{R}}$, and we write $|\underline{s}| = p + 1$. Clearly, $\underline{S'}$ is strongly normalising and $|\underline{S'}| \leqslant p$. By Lemma 5.8, $\mathsf{Supp}\left(\underline{S'}\right) \subset \mathsf{N}_{\mathsf{R}}(p)$. Then Lemma 5.3 implies $\mathsf{Supp}\left(\sigma'\right) = \mathsf{Supp}\left(\underline{S'}\right) \cup \mathsf{Supp}\left(\underline{T}\right) \subset \mathsf{N}_{\mathsf{R}}$. Hence $\|\sigma'\|$ is well defined.

We now prove that $\|\sigma'\| \prec \|\sigma\|$. The following two facts provide a sufficient condition:

(i) For all $q > |\underline{s}|$, the multiplicity of q in $\|\sigma'\|$ is the same as in $\|\sigma\|$.
(ii) The multiplicity of $|\underline{s}|$ in $\|\sigma'\|$ is stricty less than in $\|\sigma\|$.

(i) This fact boils down to the following equation

$$\sum_{|\underline{t}|} \mathsf{w}_{\underline{t}}(\sigma) = \sum_{|\underline{t}|} \mathsf{w}_{\underline{t}}\left(\sigma'\right)$$

for all $q > |\underline{s}|$. It is then sufficient to show that for $q > |\underline{s}|$ and for all $\underline{t}$ such that $|\underline{t}| = q$, we have $\mathsf{w}_{\underline{t}}\left(\sigma'\right) = \mathsf{w}_{\underline{t}}(\sigma)$. Since $\mathsf{Supp}\left(\underline{S'}\right) \subset \mathsf{N}_{\mathsf{R}}(p)$ and $p < q$, we deduce that $\underline{S'}_{(\underline{t})} = 0$ and $\sigma'_{(\underline{t})} = \underline{T}_{(\underline{t})} = \sigma_{(\underline{t})}$, and we conclude.

(ii) Similarly, for this fact we must show that

$$\sum_{|\underline{t}|=|\underline{s}|} \mathsf{w}_{\underline{t}}(\sigma) > \sum_{|\underline{t}|=|\underline{s}|} \mathsf{w}_{\underline{t}}\left(\sigma'\right).$$

Let $\underline{t}$ be such that $|\underline{t}| = |\underline{s}|$. With the same argument as above, $\underline{S'}_{(\underline{t})} = 0$, so $\sigma'_{(\underline{t})} = \underline{T}_{(\underline{t})}$. So if $t \neq s$, we have $\sigma'_{(\underline{t})} = \sigma_{(\underline{t})}$, hence $\mathsf{w}_{\underline{t}}\left(\sigma'\right) = \mathsf{w}_{\underline{t}}(\sigma)$. Moreover, by Lemma 5.1, $\mathsf{w}_{\underline{s}}(\sigma) = \mathsf{w}\left(a + \underline{T}_{(\underline{s})}\right) \geqslant \mathsf{w}(a) + \mathsf{w}_{\underline{s}}\left(\underline{T}\right)$ and $\mathsf{w}(a) > 0$. Since $\underline{T}_{(\underline{s})} = \sigma'_{(\underline{s})}$, we obtain $\mathsf{w}_{\underline{s}}(\sigma) > \mathsf{w}_{\underline{s}}\left(\sigma'\right)$. □

We can now state the final theorem of this subsection.

Theorem 5.11. The set of all strongly normalising terms is $\mathsf{R}\langle \mathsf{N}_{\mathsf{R}} \rangle$.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

Proof. One inclusion is Lemma 5.8; the other follows from Lemma 5.10 and the fact that the multiset order is a well-order. □

5.2. *Saturated sets*

We now define a notion of saturation on sets of simple terms, and then prove that N_R is saturated. Here the conditions we imposed on R are crucial since the proof relies heavily on Theorem 5.11.

Definition 5.12. Let $\mathcal{X}$ be a set of simple terms. An $\mathcal{X}$-redex is a simple term of the following shape:

$$\sigma = \underline{(\lambda x\, s)\, T}$$

where $\underline{s} \in \mathcal{X}$ and $\underline{T} \in \mathsf{R}\langle \mathcal{X} \rangle$. We write $\mathrm{Red}(\sigma)$ for the term obtained by firing this redex: $\mathrm{Red}(\sigma) = \underline{s\,[T/x]}$.

Definition 5.13. The set $\mathcal{X}$ is *saturated* if for all N_R-redex σ and all $\tau_1, \dots, \tau_n \in \mathsf{R}\langle \mathsf{N}_\mathsf{R} \rangle$, $(\mathrm{Red}(\sigma))\, \tau_1 \cdots \tau_n \in \mathsf{R}\langle \mathcal{X} \rangle$ implies $(\sigma)\, \tau_1 \cdots \tau_n \in \mathcal{X}$.

Lemma 5.14. The set N_R is saturated.

Proof. We have to prove that, for all N_R-redex σ and all $\tau_1, \dots, \tau_n \in \mathsf{R}\langle \mathsf{N}_\mathsf{R} \rangle$, if $(\mathrm{Red}(\sigma))\, \tau_1 \cdots \tau_n \in \mathsf{R}\langle \mathsf{N}_\mathsf{R} \rangle$, then $(\sigma)\, \tau_1 \cdots \tau_n \in \mathsf{N}_\mathsf{R}$. We write $\sigma = \underline{(\lambda x\, s)\, T_0}$ where $\underline{s} \in \mathsf{N}_\mathsf{R}$ and $\underline{T_0} \in \mathsf{R}\langle \mathsf{N}_\mathsf{R} \rangle$, and write $\tau_i = \underline{T_i}$ for each i. With this notation, we are led to prove that for all $\underline{s} \in \mathsf{N}_\mathsf{R}$ and all $\underline{T_0}, \dots, \underline{T_n} \in \mathsf{R}\langle \mathsf{N}_\mathsf{R} \rangle$, if

$$\underline{(s\,[T_0/x])\, T_1 \cdots T_n} \in \mathsf{R}\langle \mathsf{N}_\mathsf{R} \rangle, \tag{9}$$

then

$$\rho = \underline{(\lambda x\, s)\, T_0 \cdots T_n} \in \mathsf{N}_\mathsf{R}.$$

By Theorem 5.11, each $\underline{T_i}$ is strongly normalising. We prove the result by induction on $|\underline{s}| + \sum_{i=0}^{n} \left|\underline{T_i}\right|$. By Lemma 5.5, it is sufficient to show that for all ρ' such that $\rho \to \rho'$, we have ρ' is strongly normalising. The reduction $\rho \to \rho'$ can occur at the following positions:

- at the root of the N_R-redex;
- inside $\underline{s}$;
- inside one of the $\underline{T_i}$'s.

Head reduction In the first case, which is the only possible one if $|\underline{s}| + \sum_{i=0}^{n} \left|\underline{T_i}\right| = 0$, we have $\rho' = (\mathrm{Red}(\sigma))\, \tau_1 \cdots \tau_n$, so hypothesis (9) applies directly.

Reduction in the function Consider the case in which reduction occurs inside $\underline{s}$. So $\rho' = \underline{(\lambda x\, S')\, T_0 \cdots T_n}$ with $\underline{s} \to \underline{S'}$. We write the canonical term $S' = \sum_{l=1}^{q} a_l s'_l$ and, for all $l \in \{1, \dots, q\}$, define $\rho'_l = \underline{(\lambda x\, s'_l)\, T_0 \cdots T_n}$ so that $\rho' = \sum_{l=1}^{q} a_l \rho'_l$. It is then sufficient to prove that for all $l \in \{1, \dots, q\}$, we have $\rho'_l \in \mathsf{N}_\mathsf{R}$. For all l, we have $\left|\underline{s'_l}\right| < |\underline{s}|$ and the induction hypothesis applies to the data $\underline{s'_l}, \underline{T_0}, \dots, \underline{T_n}$. Hence it is sufficient to show that $\underline{(s'_l\,[T_0/x])\, T_1 \cdots T_n} \in \mathsf{R}\langle \mathsf{N}_\mathsf{R} \rangle$. By hypothesis (9), $(\mathrm{Red}(\sigma))\, \tau_1 \cdots \tau_n \in$

$\mathsf{R}\langle\mathsf{N_R}\rangle$. Since $\underline{s} \to \underline{S'}$, Corollary 3.13 and Lemma 3.5 imply $(\mathrm{Red}(\sigma))\,\tau_1\cdots\tau_n \;\tilde{\to}^*\; \underline{\sum_{l=1}^{q} a_l (s'_l\,[T_0/x])\,T_1\cdots T_n}$. Hence each $\underline{(s'_l\,[T_0/x])\,T_1\cdots T_n} \in \mathsf{R}\langle\mathsf{N_R}\rangle$ by Lemma 5.3.

Reduction in an argument Consider the case in which reduction occurs inside $\underline{T_i}$, so $\rho' = \underline{(\lambda x\, s)\,T_0\cdots T'_i\cdots T_n}$ with $\underline{T_i} \;\tilde{\to}\; \underline{T'_i}$. Since $\left|\underline{T'_i}\right| < \left|\underline{T_i}\right|$, the induction hypothesis applies to the data $\underline{s}, \underline{T_0}, \ldots, \underline{T'_i}, \ldots, \underline{T_n}$. Hence it is sufficient to show that (9) holds for that data, in other words,

$$\underline{(s\,[T_0/x])\,T_1\cdots T'_i\cdots T_n} \in \mathsf{R}\langle\mathsf{N_R}\rangle$$

or, if $i = 0$,

$$\underline{(s\,[T'_0/x])\,T_1\cdots T_n} \in \mathsf{R}\langle\mathsf{N_R}\rangle\,.$$

We can conclude directly, since this is a $\tilde{\to}^*$-reduct of

$$(\mathrm{Red}(\sigma))\,\tau_1\cdots\tau_n \in \mathsf{R}\langle\mathsf{N_R}\rangle$$

by contextuality of $\tilde{\to}^*$, plus Proposition 2.6 if $i = 0$. □

5.3. *Reducibility*

To each simple type, we associate a saturated subset of $\mathsf{N_R}$ as follows.

Definition 5.15. If $\mathscr{X}$ and $\mathscr{Y}$ are sets of simple terms, we define $\mathscr{X} \to \mathscr{Y} \subseteq \Delta_{\mathsf{R}}$ by

$$\mathscr{X} \to \mathscr{Y} = \{\sigma \in \Delta_{\mathsf{R}};\quad \text{for all } \tau \in \mathsf{R}\langle\mathscr{X}\rangle, (\sigma)\,\tau \in \mathscr{Y}\}\,.$$

Proposition 5.16. If $\mathscr{X}, \mathscr{X}', \mathscr{Y}, \mathscr{Y}' \subseteq \Delta_{\mathsf{R}}$ are such that $\mathscr{X} \subseteq \mathscr{X}'$ and $\mathscr{Y}' \subseteq \mathscr{Y}$, then $\mathscr{X}' \to \mathscr{Y}' \subseteq \mathscr{X} \to \mathscr{Y}$.

Lemma 5.17. If $\mathscr{S}$ is a saturated set and $\mathscr{X} \subseteq \mathsf{N_R}$, then $\mathscr{X} \to \mathscr{S}$ is saturated.

Proof. The proof is straightforward from the definitions of saturation and $\mathscr{X} \to \mathscr{S}$. □

Definition 5.18. We define the interpretation A^* of type A by induction on A:

— $\phi^* = \mathsf{N_R}$ if ϕ is a basic type;
— $(A \to B)^* = A^* \to B^*$.

Definition 5.19. Let $\mathsf{E_R}$ be the set of all simple terms σ of shape $\sigma = (\underline{x})\,\tau_1\cdots\tau_n$, where $\tau_1, \ldots, \tau_n \in \mathsf{R}\langle\mathsf{N_R}\rangle$. These are called *neutral terms.*

Lemma 5.20. The following inclusions hold:

$$\mathsf{E_R} \subseteq (\mathsf{N_R} \to \mathsf{E_R}) \subseteq (\mathsf{E_R} \to \mathsf{N_R}) \subseteq \mathsf{N_R}.$$

Proof. Of course, $\mathsf{E_R} \subseteq \mathsf{N_R}$, which gives the central inclusion by Proposition 5.16. The first inclusion holds by the definition of $\mathsf{E_R}$. If $\tau \in \mathsf{E_R} \to \mathsf{N_R}$, let x be any variable, $\underline{x} \in \mathsf{E_R}$ and we have $(\tau)\,\underline{x} \in \mathsf{N_R}$, which implies $\tau \in \mathsf{N_R}$ by Lemma 3.4, which gives the last inclusion. □

Corollary 5.21. For any type A, we have $\mathsf{E_R} \subseteq A^* \subseteq \mathsf{N_R}$.

5.4. *Adequation*

We now conclude the strong normalisation proof, that is, we prove that every simply typed term lies in the interpretation of its type. More formally, we prove the following theorem.

Theorem 5.22. Let σ be a term and assume that

$$x_1 : A_1, \dots, x_m : A_m \vdash_{\mathsf{R}} \sigma : A$$

is derivable. Let $\sigma_1 \in \mathsf{R}\langle A_1^* \rangle, \dots, \sigma_m \in \mathsf{R}\langle A_m^* \rangle$. Then

$$\sigma[\sigma_1, \dots, \sigma_m / x_1, \dots, x_m] \in \mathsf{R}\langle A^* \rangle .$$

Proof. Write $\tau = \sigma[\sigma_1, \dots, \sigma_m / x_1, \dots, x_m]$. We prove $\tau \in \mathsf{R}\langle A^* \rangle$ by induction on $\mathsf{can}(\sigma)$:

Variable $\sigma = \underline{x_i}$ for some i and $A = A_i$.
Then $\tau = \sigma_i \in \mathsf{R}\langle A_i^* \rangle$ by hypothesis.

Application $\sigma = \underline{(s)\,T}$ with $x_1 : A_1, \dots, x_m : A_m \vdash s : B \to A$ and $x_1 : A_1, \dots, x_m : A_m \vdash T : B$.
By the induction hypothesis,

$$\underline{s}[\sigma_1, \dots, \sigma_m / x_1, \dots, x_m] \in \mathsf{R}\langle (B \to A)^* \rangle$$

and

$$\underline{T}[\sigma_1, \dots, \sigma_m / x_1, \dots, x_m] \in \mathsf{R}\langle B^* \rangle .$$

Hence $\tau \in \mathsf{R}\langle A^* \rangle$ by the definition of $B^* \to A^*$.

Abstraction $\sigma = \underline{\lambda x\, s}$ and $A = B \to C$ with

$$x_1 : A_1, \dots, x_m : A_m, x : B \vdash s : C.$$

We assume x is distinct from every x_i and does not occur free in any $\mathsf{can}(\sigma_i)$. Then $\tau = \underline{\lambda x\, S'}$ with

$$\underline{S'} = \underline{s}[\sigma_1, \dots, \sigma_m / x_1, \dots, x_m].$$

We show that $\tau \in \mathsf{R}\langle (B \to C)^* \rangle$ using the definition of $B^* \to C^*$. Let $\underline{T} \in \mathsf{R}\langle B^* \rangle$. We have to prove $\underline{(\lambda x\, S')\, T} \in \mathsf{R}\langle C^* \rangle$. Since C^* is saturated, it is sufficient to show that $\underline{S'[T/x]} \in \mathsf{R}\langle C^* \rangle$. By Proposition 2.4,

$$\underline{S'[T/x]} = \underline{s}[\underline{T}, \sigma_1, \dots, \sigma_m / x, x_1, \dots, x_m]$$

and we can conclude by the induction hypothesis applied to $\underline{s}$.

Linear combinations $\sigma = \underline{\sum_{i=1}^{n} a_i s_i}$ and $\Gamma \vdash s_i : A$ for all $i \in \{1, \dots, n\}$.
Then, by the induction hypothesis, each $\underline{s_i}[\sigma_1, \dots, \sigma_m / x_1, \dots, x_m] \in \mathsf{R}\langle A^* \rangle$ and we can conclude. □

We have the following corollary of Theorem 5.22.

Theorem 5.23. All weakly typable terms are strongly normalising.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

Proof. Let $\sigma \in \mathsf{R}\langle\Delta_\mathsf{R}\rangle$ be such that $x_1 : A_1, \ldots, x_m : A_m \vdash_\mathsf{R} \sigma : A$ is derivable. For all $i \in \{1, \ldots, n\}$, since $\mathsf{E}_\mathsf{R} \subseteq A_i^*$, we have $\underline{x_i} \in \mathsf{R}\langle A_i^*\rangle$. Hence

$$\sigma = \sigma\,[\underline{x_1}, \ldots, \underline{x_m}/x_1, \ldots, x_m] \in \mathsf{R}\langle A^*\rangle$$

by Theorem 5.22, and we can conclude by Corollary 5.21 and Theorem 5.11. □

5.5. *Weak normalisation scheme*

Remember that we forced strong conditions on R at the beginning of this section. One can get rid of this restriction by slightly changing the notion of normal form, as has already been noted in Ehrhard and Regnier (2003). In the following, we provide a full development of their argument.

Definition 5.24. We define *pre-normal terms* and *pre-neutral terms* by the following inductive statements:

— $\sigma \in \Delta_\mathsf{R}$ is a pre-neutral term if $\sigma = \underline{x}$ with $x \in \mathscr{V}$, or $\sigma = \underline{(s)\,T}$, where $\underline{s}$ is a pre-neutral term and $\underline{T}$ is a pre-normal term;
— $\sigma \in \Delta_\mathsf{R}$ is a simple pre-normal term if σ is pre-neutral, or $\sigma = \underline{\lambda x\, s}$ where $\underline{s}$ is a simple pre-normal term;
— σ is a pre-normal term if, for all $\underline{s} \in \mathsf{Supp}(\sigma)$, $\underline{s}$ is a simple pre-normal term.

Intuitively, pre-normal terms are those terms σ such that $\mathsf{can}(\sigma)$ contains no redex. Hence, we have the following proposition.

Proposition 5.25. If R is positive, pre-normal terms are exactly normal terms (and pre-neutral terms are exactly neutral terms).

A rig of polynomials Let R be any rig and Ξ be a set of variables in bijection with R:

— to every $a \in \mathsf{R}$ we associate $\xi_a \in \Xi$ such that $\xi_a = \xi_b$ if and only if $a = b$; and
— $\Xi = \{\xi_a;\ a \in \mathsf{R}\}$.

Definition 5.26. Let $\mathsf{P} = \mathbf{N}[\Xi]$ be the rig of polynomials with non-negative integer coefficients over variables in Ξ. If $P \in \mathsf{P}$ and $f : \mathsf{R} \longrightarrow \mathsf{R}'$ where R' is any rig, we use

$$P\{a \mapsto f(a)\}$$

to denote the valuation of P at f, that is, the scalar (in R') obtained by replacing each ξ_a in P by $f(a)$ for all $a \in \mathsf{R}$.

Definition 5.27. If $P \in \mathsf{P}$, we use $[\![P]\!]$ to denote the value of P in R:

$$[\![P]\!] = P\{a \mapsto a\} \in \mathsf{R}.$$

Lemma 5.28. The rig P is finitely splitting and has no zero divisor.

Proof. The width function is exactly the sum of all coefficients:

$$\mathsf{w}(P) = P\{a \mapsto 1\} \in \mathbf{N}.$$ □

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

Hence Theorem 5.23 applies and we obtain the following corollary.

Corollary 5.29. All weakly typable terms in $\mathsf{P}\langle\Delta_{\mathsf{P}}\rangle$ are strongly normalising.

We now extend the valuation of a term in $\mathsf{P}\langle\Delta_{\mathsf{P}}\rangle$ as the term in $\mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ obtained by replacing each polynomial coefficient with its value.

Definition 5.30. We define $[\![\cdot]\!] : \mathsf{P}\langle\Delta_{\mathsf{P}}\rangle \longrightarrow \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ by induction on terms:

$$\begin{aligned} [\![\underline{x}]\!] &= x \\ [\![\underline{\lambda x\, s}]\!] &= \lambda x\, [\![\underline{s}]\!] \\ [\![\underline{(s)\, T}]\!] &= ([\![\underline{s}]\!])\, [\![\underline{T}]\!] \\ \left[\!\!\left[\underline{\sum_{i=1}^{n} P_i s_i}\right]\!\!\right] &= \sum_{i=1}^{n} [\![P_i]\!]\, [\![\underline{s_i}]\!] \,. \end{aligned}$$

Proposition 5.31. For all $\sigma \in \mathsf{P}\langle\Delta_{\mathsf{P}}\rangle$, if σ is a pre-normal term, then $[\![\sigma]\!] \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ is a pre-normal term.

Lemma 5.32. For all $\sigma, \sigma' \in \mathsf{P}\langle\Delta_{\mathsf{P}}\rangle$, if $\sigma \stackrel{\sim}{\to} \sigma'$, then $[\![\sigma]\!] \stackrel{\sim}{\to}^{*} [\![\sigma]\!]'$.

Proof. The proof is easy by induction on reduction $\sigma \stackrel{\sim}{\to} \sigma'$. □

Definition 5.33. For all $M \in \Lambda_{\mathsf{R}}$, we define $\check{M} \in \Lambda_{\mathsf{P}}$ as the permutative term obtained from M by replacing every coefficient a with the monomial χ_a.

Lemma 5.34. For all $\underline{S} \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$, we have $\underline{S} = \left[\!\!\left[\underline{\check{S}}\right]\!\!\right]$.

Proof. For all $\underline{s} \in \mathsf{Supp}(\underline{S})$, we have $\underline{S}_{(\underline{s})} = S_{(s)} = [\![\xi_{S_{(s)}}]\!] = \left[\!\!\left[\check{S}_{(\check{s})}\right]\!\!\right]$. □

Lemma 5.35. Let $\underline{S} \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$. If $\Gamma \vdash_{\mathsf{R}} \underline{S} : A$, then $\Gamma \vdash \check{S} : A$.

Proof. It is easy to prove by induction on the permutative term M that if $\Gamma \vdash M : A$, then $\Gamma \vdash \check{M} : A$. □

Theorem 5.36. Let $\sigma \in \mathsf{R}\langle\Delta_{\mathsf{R}}\rangle$ be a weakly typable term. Then σ is weakly normalising in the sense that it reduces to a pre-normal form.

Proof. If σ is weakly typable, then, by Lemma 5.35, $\check{\sigma}$ is typable. By Theorem 5.23, $\check{\sigma}$ is strongly normalising, so $\check{\sigma} \stackrel{\sim}{\to}^{*} \tau$ where τ is normal. By Proposition 5.25, τ is pre-normal, and, by Proposition 5.31, so is $[\![\tau]\!]$. By Lemma 5.32, $\sigma \stackrel{\sim}{\to}^{*} [\![\tau]\!]$, which gives the conclusion. □

Recall that if R is positive, then every pre-normal form is a normal form, and in this case, Theorem 5.36 states a genuine weak normalisation property.

6. Other approaches and related work

Undeterminate Forms

It is noteworthy that the collapse we described in Section 3.3 involves a term ∞_σ such that $\infty_\sigma \mathrel{\tilde{\rightarrow}^*} n\sigma + \infty_\sigma$, for all $n \in \mathbf{N}$, so reduction of ∞_σ generates an unbounded amount of σ. This is not a surprise, since the untyped algebraic λ-calculus involves both linear algebra and arbitrary fixed points. The term $\infty_\sigma + (-1)\infty_\sigma$ is then analoguous to the well-known indeterminate form $\infty - \infty$ of the affinely extended real number line (that is $\mathbf{R} \cup \{-\infty, \infty\}$, the two-point compactification of $\mathbf{R}$, where the usual operations can only be partially extended). The collapse of reduction in the presence of negative scalars follows from the fact that we consider $\infty_\sigma - \infty_\sigma = \underline{\mathbf{0}}$.

Notice that our observations do not depend on equations (1) and (2). As a matter of fact, if there exists $\eta \in \mathsf{R}$ with $1 + \eta = 0$, then any contextual equivalence relation $\cong$ defined on raw terms such that

— $\cong$ contains β-reduction, that is, $(\lambda x\, M)\, N \cong M\,[N/x]$ for all $M, N \in \Lambda_{\mathsf{R}}$,

— $\cong$ contains R-module equations (groups of equations (6) and (7))

is unsound. Indeed, we can define $\infty_M \in \Lambda_{\mathsf{R}}$ for all $M \in \Lambda_{\mathsf{R}}$, and then $\infty_M + \eta\infty_M$ is $\cong$-equal to both M and $\mathbf{0}$:

$$\infty_M + \eta\infty_M \cong (1+\eta)\infty_M \cong \mathbf{0}$$

and

$$\begin{aligned} \infty_M + \eta\infty_M &\cong (M + \infty_M) + \eta\infty_M \quad \text{by iterated } \beta\text{-reductions} \\ &\cong M + (\infty_M + \eta\infty_M) \\ &\cong M + (1+\eta)\infty_M \\ &\cong M + \mathbf{0} \\ &\cong M. \end{aligned}$$

One seemingly natural variant of one-step reduction is the following one, which we outlined in the introduction. Instead of (4), we extend reduction from simple terms to all terms using

$$\sigma \mathrel{\hat{\rightarrow}} \sigma' \text{ if } \sigma = \underline{as + T} \text{ and } \sigma' = \underline{aS' + T}, \text{ with } a \neq 0,\ T_{(s)} = 0 \text{ and } \underline{s} \rightarrow \underline{S'}. \tag{10}$$

As far as reduction is concerned, this amounts to restricting the syntax to canonical forms of terms. Note that this is not contextual in the sense of Definition 2.5. This is still unsound in general, however, since we can reproduce the argument of Section 3.3.2, but replacing $a\infty_\sigma + b\infty_\sigma$ with $a\infty_\sigma + b(\lambda x\, x)\,\infty_\sigma$.

We have already mentioned another technique to deactivate coefficients and tame $\triangleq$ during reduction by replacing the coefficients of a term with formal variables, then reducing some steps, and, finally, replacing the variables with their values. Reduction $\hat{\rightarrow}$ can be seen as a strategy in this setting. In particular, $\hat{\rightarrow}$ is well behaved as far as normalisation is concerned since the trick involving rational coefficients is no longer possible, and (weakly) typed terms are strongly normalising.

https://doi.org/10.1017/S0960129509990089 Published online by Cambridge University Press

A possible fix for the collapse while retaining the algebraic structure of the calculus might involve typing in order to ward off arbitrary fixed points. Then one has to introduce some typed notion of reduction, though we have seen that typability is not even preserved under our notion of reduction. This is the subject of current work in connection with the quantitative semantics of simply typed ordinary λ-calculus in the finiteness spaces of Ehrhard (2005).

Algebraic rewriting

In Arrighi and Dowek (2008), the authors introduced the linear algebraic λ-calculus. The background setting is quite unrelated to ours since their work is designed to provide a framework for quantum computation. In particular, terms represent linear operators, so application is bilinear rather than linear in the function only. In addition to this distinction, their approach to λ-calculus with linear combinations of terms can be contrasted with ours in other ways: they consider terms up to $\equiv$ rather than some variant of $\triangleq$, and handle the identities between linear combinations, together with analogues of (1) and (2), as reduction rules.

However, confronted with problems similar to those we exposed above due to the presence of negative coefficients, they opted for a completely different solution, which is far more natural in their setting. Their solution is to restrict those reduction rules involving rewriting of linear combinations to closed terms in normal form. This allows them to tame some of the intrinsic potential infinities of the pure λ-calculus, and avoid having to consider indeterminate forms. They then prove confluence for the whole system with their restrictions.

This opens up interesting perspectives for future work, which is already the subject of a collaboration with Arrighi and Dowek. In particular, it seems that a system similar to that of Arrighi and Dowek (2008) can be designed in the setting of the algebraic λ-calculus. Moreover, one can view the divergence in the treatment of linearity in each work as a manifestation of the call-by-name (CBN) *vs.* call-by-value (CBV) duality: Arrighi–Dowek's linear algebraic λ-calculus is intrinsically a CBV system, while our algebraic λ-calculus is rooted in the CBN translation of λ-calculus in linear logic (recall that it originated from the presentation of Ehrhard and Regnier's differential λ-calculus). It is a matter of particular interest to determine whether the calculi enjoy the same relationship with each other as that known to exist between the CBN and CBV flavours of pure λ-calculus.

Acknowledgements

I am deeply indebted to René David, who suggested the use of the multiset order in the proof of Theorem 5.11.

I also wish to thank the anonymous referee for her/his careful reading and many useful suggestions and comments concerning both style and mathematical content.

I have also enjoyed enlightening discussions with Pablo Arrighi and Gilles Dowek.

References

Arrighi, P. and Dowek, G. (2005) A computational definition of the notion of vectorial space. *Electr. Notes Theor. Comput. Sci.* **117** 249–261.

Arrighi, P. and Dowek, G. (2008) Linear-algebraic lambda-calculus: higher-order, encodings, and confluence. In: Voronkov, A. (ed.) RTA. *Springer-Verlag Lecture Notes in Computer Science* **5117** 17–31.

Ehrhard, T. (2001) On Köthe sequence spaces and linear logic. *Mathematical Structures in Computer Science* **12** 579–623.

Ehrhard, T. (2005) Finiteness spaces. *Mathematical Structures in Computer Science* **15** (4) 615–646.

Ehrhard, T. and Regnier, L. (2003) The differential lambda-calculus. *Theoretical Computer Science* **309** 1–41.

Girard, J.-Y. (1987) Linear logic. *Theoretical Computer Science* **50** 1–102.

Girard, J.-Y. (1988) Normal functors, power series and lambda-calculus. *Annals of Pure and Applied Logic* **37** (2) 129–177.

Krivine, J.-L. (1990) *Lambda-calcul, types et modèles*, Masson, Paris.

Peterson, G. E. and Stickel, M. E. (1981) Complete sets of reductions for some equational theories. *J. ACM* **28** (2) 233–264.

Vaux, L. (2007a) The differential $\lambda\mu$-calculus. *Theoretical Computer Science* **379** (1–2) 166–209.

Vaux, L. (2007b) On linear combinations of λ-terms. In: Baader, F. (ed.) RTA. *Springer-Verlag Lecture Notes in Computer Science* **4533** 374–388.