Syntax-Free Developments

Vincent van Oostrom
Universität Innsbruck, Austria

Abstract
We present the Z-property and instantiate it to various rewrite systems: associativity, positive braids, self-distributivity, the lambda-calculus, lambda-calculi with explicit substitutions, orthogonal TRSs, .... The Z-property is proven equivalent to Takahashi’s angle property by means of a syntax-free notion of development. We show that several classical consequences of having developments such as confluence, normalisation, and recurrence, can be regained in a syntax-free way, and investigate how the notion corresponds to the classical syntactic notion of development in term rewriting.

2012 ACM Subject Classification Theory of computation → Equational logic and rewriting

Keywords and phrases rewrite system, confluence, normalisation, recurrence

Digital Object Identifier 10.4230/LIPIcs.FSCD.2021.24

Acknowledgements Patrick Dehornoy introduced me to the main themes presented here, and indeed this paper was always intended to be a joint one. His work continues to be an inspiration. I want to thank Bertram Felgenhauer, Julian Nagele, and Christian Sternagel for discussions on their Isabelle formalisations of the Z-property.

Dedicated to Patrick Dehornoy

1 Introduction

Confluence of rewrite systems is discussed in order-theoretic terms on the first page of [25]. It expresses the existence of an upper bound\(^1\) for pairs of objects having a common lower bound, in the quasi-order obtained by the reflexive–transitive closure of a rewrite system. Qualifying confluence proof-methods from this order-theoretic perspective, Newman’s Lemma is seen to construct the greatest upper bound (the normal form) and the Tait–Martin-Löf (TML) method [4] the least upper bound [21, 38].\(^2\) The Z-property, depicted in Fig. 1 and formally defined in the preliminaries, introduced here is based on constructing an upper bound for sets of objects having a common single-step lower bound. The choice of upper bound is arbitrary but should be monotonic: increasing the single-step lower-bound should increase the constructed upper bound. In complexity, establishing some upper bound is often much shorter and simpler than getting a tight upper bound. The choice offered by the Z-property enables the same for proving confluence, as we illustrate in Sect. 3.

Skolemising the existence of upper bounds gives rise to a function •\(^3\) mapping each object \(a\) to the chosen upper bound \(a^*\) of objects \(b\) such that \(a \rightarrow b\), i.e. having \(a\) as single-step lower bound. Accordingly, we define the many-step rewrite strategy \(\leadsto\) to rewrite \(a\) into \(a^*\). For instance, taking as upper bound of a term \(t\) the term \(t^*\) obtained by a complete development of the full set of redexes in \(t\), \(\leadsto\) is known as the Gross–Knuth/full substitution strategy in the \(\lambda\)-calculus/term rewriting [4, 38]. Based on •, the classical notion of a

---

\(^1\) [25] employs the reverse order, so speaks of existence of lower bounds.

\(^2\) Newman leaves studying least upper bounds for later [25, p. 223] but we didn’t find later work by him on this. TML in fact gives least upper bounds only up to permutation equivalence [21, 38].

\(^3\) We will speak of the bullet function with the suggestion \(\leadsto\) is bullet-fast; cf. Sect. 4.1.
development [5, 4, 38] can be given a syntax-free definition as \( a \rightarrow b \) if \( a \rightarrow b \rightarrow a^* \); that is, \( a \) develops to \( b \) if \( b \) is between \( a \) and \( a^* \); with our notations suggesting that \( \rightarrow \) is a development that is not as full as \( \bullet \rightarrow \) is. In Sect. 4 we first show that if the Z-property holds then several results (on confluence, normalisation, and recurrence) can be obtained in a syntax-free way, i.e. in terms of \( \bullet \rightarrow \) and \( \rightarrow \). Next we investigate for term rewrite systems in how far our syntax-free definition of developments corresponds or can be made to correspond to the traditional syntactic definition, and show they correspond in the absence of syntactic accidents.

**Remark 1.** Thinking of reduction steps and reductions to normal form as small respectively big step semantics, \( \bullet \rightarrow \) can be seen as a medium step semantics; although \( \bullet \rightarrow \)-steps need not directly yield a normal form, they are monotonic. This may be suitable in a setting where for a step \( a \rightarrow b \), the semantics of \( b \) should be greater than that of \( a \), i.e. approximate better.

## 2 Preliminaries

We define our key notions for abstract rewriting with which we assume basic familiarity [38].

**Definition 2.** A rewrite system is a system comprising a set of objects, a set of (rewrite) steps, and functions \( \text{src}, \text{tgt} \) mapping a step to its source, target object. Two steps are called co-initial if they have the same sources, co-final if they have the same targets, and composable if the target of the former is the source of the latter. The corresponding pair of steps is then called, respectively, a peak, a valley, and consecutive.

**Remark 3.** We follow [25] in taking steps as first-class citizens of rewrite systems and speak of a rewrite relation (only) if there is at most one step between any two objects.

We use arrow-like notations to denote rewrite systems and their steps, let \( a, b, \ldots \) range over objects, and \( \phi, \psi, \ldots \) over steps. Sources and targets naturally extend to peaks, valleys, and consecutive steps; e.g., the source of a peak is the common source of its steps and its target is its pair of targets.

**Definition 4.** A rewrite system \( \rightarrow \) has the (see Fig. 1):
- diamond property if for every peak there is a composable valley;
- angle property if there is map \( \bullet \) such that \( b \rightarrow a^* \) for every \( a \) and step \( a \rightarrow b \); and
- Z property if there is a map \( \bullet \) such that \( b \rightarrow a^* \rightarrow b^* \) for every \( a \) and step \( a \rightarrow b \).

where \( \rightarrow \) denotes reduction, finite (possibly empty) composition of steps. A map \( \bullet \) is extensive if \( a \rightarrow a^* \) for all \( a \), and induces a rewrite system \( \bullet \rightarrow \) having the same objects as \( \rightarrow \) and steps \( a \rightarrow a^* \) for all \( a \) not in \( \rightarrow \)-normal form.

**Remark 5.** The diamond and angle properties are relatively standard in rewriting, see e.g. [38, Def. 1.1.8]; our angle property is the Skolemisation of the triangle property there. We obtained the Z-property in 2007 by abstracting Dehornoy’s proof-method for showing confluence of self-distributivity [6] with preliminary results distributed and presented at
Two angles make a diamond, but the angle property is stronger than the diamond property. If the Z-property holds \( \bullet \) is monotonic on reductions: if \( a \rightarrow b \) then \( a^* \rightarrow b^* \) (by induction).

**Example 6.** Less-than \(<\) on \( \mathbb{Z} \) has the diamond but not the angle property for lack of upper bounds of infinite sets of numbers. Note that the predecessor relation on \( \mathbb{Z} \) does have the angle property, despite inducing the same quasi-order as \(<\).

The following simple but key result was the starting point of our investigations on the Z-property. It hinges on a syntax-free definition of the classical notion of development [4, 38].

**Definition 7.** For rewrite system \( \rightarrow \) and map \( \bullet \) on its objects, the \( \bullet \)-development rewrite system \( \rightarrow \bullet \) has the objects of \( \rightarrow \) and a step \( a \rightarrow \bullet b \) for each pair of \( \rightarrow \)-reductions \( a \rightarrow b \rightarrow a^* \).

One may think of \( b \) as being between \( a \) and \( a^* \) and of \( \rightarrow \bullet \) as comprising prefixes or left-divisors \( \rightarrow \) w.r.t. composition (for sources not in normal form).

**Lemma 8.** Let \( \rightarrow \) be a rewrite system.

1. \( \rightarrow \) has the Z-property iff some \( \rightarrow' \) such that \( \rightarrow \subseteq \rightarrow' \subseteq \rightarrow \) has the angle property;\(^4\)
2. if \( \rightarrow \) has the Z-property for \( \bullet \), then it has the Z-property for some extensive \( \star \); and
3. \( \rightarrow \) has the Z-property for an extensive \( \bullet \) iff some rewrite system \( \rightarrow' \) such that \( \rightarrow \subseteq \rightarrow' \subseteq \rightarrow \) has the angle property and \( a \rightarrow \bullet a^* \) for all \( a \).

**Proof.** We only provide a detailed proof of the first, main, item.

we show both directions taking the same bullet function \( \bullet \).

For the if-direction, assume \( \rightarrow \rightarrow' \) has the angle property, \( \rightarrow \subseteq \rightarrow' \subseteq \rightarrow \), and suppose \( a \rightarrow b \). Then by \( \rightarrow \subseteq \rightarrow' \) and the angle property for \( a \rightarrow' b \) we have \( b \rightarrow a^* \), hence \( a^* \rightarrow' b^* \) by applying the angle property again. Two angles make a Z; using \( \rightarrow' \subseteq \rightarrow \) twice, we conclude to \( b \rightarrow a^* \) and \( a^* \rightarrow b^* \).

For the only-if-direction, assume \( \rightarrow \) has the Z-property. Consider the \( \bullet \)-development rewrite system \( \rightarrow \bullet \). To show \( \rightarrow \bullet \) has the angle property, suppose \( a \rightarrow b \). By definition \( a \rightarrow b \rightarrow a^* \). Combining \( b \rightarrow a^* \) with \( a^* \rightarrow b^* \), which follows from \( a \rightarrow b \) by monotonicity of \( \bullet \), yields \( b \rightarrow a^* \) by definition of \( \rightarrow \bullet \), showing the angle property. That the first inclusion in \( \rightarrow \subseteq \rightarrow \bullet \subseteq \rightarrow \) holds follows from that \( a \rightarrow b \) entails \( b \rightarrow a^* \) by the Z-property hence by definition \( a \rightarrow b \rightarrow a^* \), and that the second inclusion holds from that \( a \rightarrow b \rightarrow a^* \) unfolds to \( a \rightarrow b \rightarrow a^* \).

one checks that defining \( \star \) to be \( \bullet \) updated to map each object that is not the source of some step to itself, works; and

one checks the additional conditions on either side in the first item. The if-direction is trivial since \( \rightarrow' \subseteq \rightarrow \) by assumption.

Adjoining being extensive to the angle property in Fig. 1 gives rise to a triangle, i.e. the second and third items reconcile both names of the property.

Although the intuition is that \( \bullet \)-developments correspond to developments, the former, by being defined in a syntax-free way, are more liberal (we will look into this in Sect. 4.4) as shown by:

\[^4\] The inclusions are relation inclusions, i.e. concern the rewrite relation underlying the rewrite systems.
Example 9. The rewrite system \( a_i \rightarrow a_{i+1} \mod 4 \) has the Z-property for the function • mapping \( a_i \) to \( a_{i+1} \mod 4 \) because \( \rightarrow \) is deterministic. Classically there are only two developments from \( a_0 \) namely to itself, the empty development, and to \( a_1 \). However, because \( \rightarrow \) is cyclic there are more •-developments, e.g. \( a_0 \rightarrow a_2 \) (since \( a_0 \rightarrow a_2 \rightarrow a_1 = a_0^* \)).

3 Examples of the Z-property

We present (non-)examples of rewrite systems having the Z-property with a focus on the diversity of the examples and the similarity of the proofs. We give proofs in as far as they could serve as blue-prints of proofs of the Z-property for related calculi. We proceed from abstract to more concrete rewrite systems.

3.1 Abstract

We investigate for some known confluence criteria for (abstract) rewrite systems [3, 38] whether or not they entail the Z-property. We assume \( \rightarrow \) is a rewrite system. In the previous section we have already seen a characterisation of the Z-property via the angle property. That the Z-property holds for deterministic (if \( a \rightarrow b \) and \( a \rightarrow c \), then \( a = b \)) systems by mapping to the next object was exemplified in Ex. 9.

Lemma 10. If \( \rightarrow \) is deterministic, then it has the Z-property.

In case a rewrite system is terminating mapping to the greatest object works.

Lemma 11. If \( \rightarrow \) is terminating, then \( \rightarrow \) has the Z-property iff \( \rightarrow \) is locally confluent.

Proof. Suppose \( \rightarrow \) is locally confluent and terminating. Let \( \bullet \) be the normal form function mapping each object to its \( \rightarrow \)-normal form. This is well-defined: the normal form exists by termination and is unique as local confluence entails confluence by Newman’s Lemma. Thus we conclude to the Z-property since if \( a \rightarrow b \) then \( b \Rightarrow a^* = b^* \). Vice versa, if \( \rightarrow \) has the Z-property for \( \bullet \) then \( a^* \) is a common reduct to all \( b \) such that \( a \rightarrow b \).

Ex. 6 shows there are confluent rewrite systems \( \rightarrow \) that do not have the Z-property but admit it in that there is a rewrite system \( \rightarrow' \) presenting the same quasi-order, i.e. \( \rightarrow = \rightarrow' \), that does have the Z-property: \(<\) does not have the Z-property but admits it as it is the reflexive–transitive closure of the predecessor relation that does have the Z-property (by being deterministic).\(^5\) By the first item of Lem. 8 a rewrite system admits the Z-property iff it admits the angle property, using for the only–if-direction that \( \rightarrow \subseteq \rightarrow' \subseteq \rightarrow \) entails \( \rightarrow \) and \( \rightarrow' \) present the same quasi-order. But there are confluent rewrite systems not admitting either.

Example 12. Consider the confluent rewrite system\(^6\) given by \( a \rightarrow b_i \rightarrow c_i \rightarrow c_{i+1} \) for \( i \in \mathbb{N} \), and suppose \( \rightarrow' \) were some presentation of it having the Z-property. Observe that then \( a \rightarrow' b_i \) for \( i \in \mathbb{N} \), since there are no objects between \( a \) and \( b_i \) in \( \rightarrow \), but there is no common upper bound to all \( b_i \) in \( \rightarrow \), so neither there is one in \( \rightarrow' \).

Remark 13. Bullet functions for the Z-property may be incomparable (comparing their images bulletwise by \( \rightarrow \)), but are preserved under composition allowing arbitrary speed-up.

---

\(^5\) If a rewrite system has the Z-property, then so does its so-called transitive reduction, but not necessarily the other way around. However note that \(<\) admits the Z-property even on \( \mathbb{R} \), e.g. by restricting to pairs of reals having distance at most 1, despite that \(<\) then has no transitive reduction.

\(^6\) The rewrite system is a variation on the rewrite systems visualised in [12, Fig. 2].
3.2 Positive braids

Positive braids have the Z-property [6] or equivalently the angle property [34],[38, Sect. 8.9].

Definition 14. The rewrite system $\mathcal{B}^+$ of (positive) braids on $\ell$ strands has:

- as objects braids, words over the Artin generators $\sigma_i$ for $1 \leq i < \ell$, modulo
  \[
  \sigma_i \sigma_j \sigma_i = \sigma_j \sigma_i \sigma_j \quad \text{if } |i - j| = 1 \tag{1}
  \]
  \[
  \sigma_i \sigma_j = \sigma_j \sigma_i \quad \text{if } |i - j| > 1 \tag{2}
  \]

- steps $w \rightarrow w \sigma_i$ for any braid $w$ and $1 \leq i < \ell$.

The equivalence generated by (1) and (2) is denoted by $\equiv$. The rewrite system $\mathcal{B}^+$ is locally confluent as illustrated in Figure 2: any pair of distinct generators $\sigma_i, \sigma_j$ either is too far apart (2) like $\sigma_1$ and $\sigma_3$ on the left, or too close together (1) like $\sigma_1$ and $\sigma_2$ on the right. See Figure 3 for two words representing the same positive braid on 6 strands. Extending a braid by a full swap, crossing all strands over another as represented by the Garside word, works, the intuition being that is the least way to extend all single steps. The proof is short and by straightforward inductions.

Lemma 15. $\mathcal{B}^+$ has the Z-property for the map that suffixes the Garside word.

Proof. The bullet function suffixing the Garside word is formally defined by $w^* := wG_\ell$, where, starting crossing from the left, the Garside word may be inductively defined by $G_0 := \varepsilon$ and if $n > 0$, then $G_n := G_{n-1} \sigma_{(n,1)}$ with $\sigma_{(i,j)} := \sigma_{i-1} \ldots \sigma_j$ crossing the $i$th strand over $i - j$ strands to its left. The key property of $G_\ell$ is that it is a so-called Garside element as each generator is both a left and right divisor of it. More specifically, we claim that for all $1 \leq i < n$ there exists a braid $G_i^\dagger$ such that (cf. Ex. 16)

\[
\sigma_i G_n \equiv G_n \equiv G_n^\dagger \sigma_n \sigma_{n-i} \quad \tag{3}
\]

From the claim we conclude to the Z-property, since for a step $w \rightarrow w \sigma_i$ then $w \sigma_i \rightarrow w \sigma_i G_n \equiv wG_n \rightarrow wG_n \sigma_{n-i} \equiv w \sigma_i G_n^\dagger \sigma_{n-i} \equiv w \sigma_i G_n$. 

Figure 2 Local confluence diagrams for positive braids.

Figure 3 Isotopic braids, $\sigma_3 \sigma_5 \sigma_1 \sigma_4 \sigma_3 \equiv \sigma_5 \sigma_4 \sigma_3 \sigma_4 \sigma_1 \sigma_3$, deformable into one another.
It remains to prove the claim (a well-known fact). The intuition for $G'_n$ is that it is the residual of $G_n$ after $\sigma_i$, i.e. what remains to be done of a full swap after swapping $i$. Formally, it may be inductively defined by $G_{n-1} := G_{n-1}\sigma_{(n,2)}$ and $G'_n := G'_{n-1}\sigma_{(n,1)}$ otherwise. Accordingly, we show (3) by induction on $n$, with trivial base case, and cases on whether or not $i = n - 1$:

$$\sigma_i G'_n = \sigma_i G_{n-1}\sigma_{(n,1)} \quad \sigma_{n-1} G'_n = \sigma_{n-1} G_{n-2}\sigma_{(n-1,1)}\sigma_{(n,2)}$$

$\equiv_{IH} G_{n-1}\sigma_{(n,1)} = \sigma_{n-1} G_{n-2}\sigma_{(n-1,1)}\sigma_{(n,2)}$

$\equiv_{IH} G'_n = G_{n-1}\sigma_{n-1-i}\sigma_{(n,1)} \quad \equiv_{(i)} G_{n-1}\sigma_{n-1-i}\sigma_{(n,1)}$

$= G'_n\sigma_{n-1-i} \quad = G_{n-1}\sigma_{(n,1)}$

where $(i)$ follows by $(2)$. $\sigma_{n-1}$ and $G_{n-2}$ commute, i.e. $\sigma_{n-1} G_{n-2} \equiv G_{n-2} \sigma_{n-1}$, as their generators are too far apart, $(ii)$ holds since for all $i + 1 < k \geq j$:

$$\sigma_k \sigma_{(i,j)} \equiv (2) \quad \sigma_{i,k+2}\sigma_k\sigma_{k+1}\sigma_{k,j} \quad \equiv (1) \quad \sigma_{i,k+2}\sigma_{k+1}\sigma_{k,j} \quad \sigma_{(i,j)}\sigma_{k+1}$$

and $(iii)$ follows from $(ii)$ by induction on $\sigma_{(n,2)}$.

**Example 16.** To see that (3) holds for $i := 2$ and $n := 4$, we first compute $G'_2 := \sigma_1\sigma_2\sigma_3\sigma_2\sigma_1$ and $G_4 := \sigma_1\sigma_2\sigma_3\sigma_2\sigma_1$, and then verify $\sigma_2\sigma_1\sigma_2\sigma_3\sigma_2\sigma_1 \equiv_1 \sigma_1\sigma_2\sigma_1\sigma_2\sigma_2 \equiv_2 \sigma_1\sigma_2\sigma_3\sigma_1\sigma_2 \equiv_1 \sigma_1\sigma_2\sigma_3\sigma_2\sigma_1\sigma_2$.

### 3.3 First-order terms

In this section we consider TRSs, i.e. first-order term rewrite systems [3, 38]. We show the Z-property holds for orthogonal TRSs for the full development and the full superdevelopment functions, for weakly orthogonal TRSs by the maximal multistep map, for associativity by an inductive normal form function, and extending that, for self-distributivity by the full distribution function. Our presentation suggests the commonality between the proofs the Z-property holds. We assume $T$ is a TRS and $\rightarrow_T$ or simply $\rightarrow$ to be its underlying rewrite system on terms $t, s, r, \ldots$. Each bullet function $\bullet$ on terms defined below is assumed to be pointwise extended to vectors of terms $\vec{t}, \vec{s}, \ldots$ and substitutions $\sigma, \tau, \ldots$. We first observe that as a corollary to Lem. 11 and Huet’s Critical Pair Lemma we immediately have:

**Corollary 17.** A terminating TRS has the Z-property iff all its critical pairs are joinable.

### 3.3.1 Orthogonal

We show orthogonal TRSs, i.e. left-linear and non-overlapping, have the Z-property.

**Example 18.** The classical example of an orthogonal TRS is Combinatory Logic (CL). It has a binary symbol $@$ and constants $K, S, I$ and rules, written in full on the left and applicatively [38, Sect. 3.3.5] on the right (making $@$ implicit, infix, and associate to the left):

$$\begin{align*}
@\langle I, x \rangle & \rightarrow x \\
@\langle @\langle K, x \rangle, y \rangle & \rightarrow x \\
@\langle @\langle @\langle S, x \rangle, y \rangle, z \rangle & \rightarrow @\langle @\langle x, z \rangle, @\langle y, z \rangle \rangle
\end{align*}$$

$Ix \rightarrow x$  
$Kxy \rightarrow x$  
$Sxyz \rightarrow xz(yz)$

For orthogonal TRSs mapping a term to the result of contracting all redexes works, the intuition being again that it is the least way of extending all single steps. This amounts to an inductive definition of the full substitution or maximal multistep strategy [38, Def. 9.3.18].
\textbf{Definition 19.} For an orthogonal TRS, \textit{full development} \( \bullet \) is inductively defined by 
\[ x^{\bullet} := x \]
\[ f(\overline{t})^{\bullet} := r^{\sigma} \quad \text{if} \quad f(\overline{t}) \text{ is a redex and } f(\overline{t}) = \ell^{r} \text{ for some } \ell \rightarrow r \text{ and substitution } \sigma \]
\[ := f(\overline{t}) \quad \text{otherwise} \]

\textbf{Example 20.} In CL, \((I(x))^{\bullet} = x\) and \((IIx)^{\bullet} = Ix\) contracting \(II\) but not the created \(Ix\).

\textbf{Remark 21.} By orthogonality, if for some redex \(t\) there is a reduction without head-steps \(t \rightarrow \ell^{r}\) for lhs of a rule \(\ell\) and substitution \(\tau\), then \(t = \ell^{r}\) for some substitution \(\sigma\) such that \(\sigma \rightarrow \tau\). Vice versa, if we have such reduction \(\ell^{r} \rightarrow t\) for some term \(t\), then \(t = \ell^{r}\) and \(\tau \rightarrow \sigma\).

\textbf{Lemma 22.}
\begin{itemize}
  \item \textit{(Extensive)} \(t \rightarrow t^{\bullet}\) for all terms \(t\);
  \item \textit{(Rhs)} \(t^{(\sigma^{\bullet})} \rightarrow (t^{\sigma})^{\bullet}\) for terms \(t\), substitutions \(\sigma\); \(t^{(\sigma^{\bullet})} = (t^{\sigma})^{\bullet}\) if \(t\) is a proper subterm of a lhs;
  \item \textit{(Z)} \(\rightarrow\) has the Z-property for the full development function.
\end{itemize}

\textbf{Proof.}

(Extensive) By induction on \(t\). If \(t\) is a variable \(x\), then \(t^{\bullet} = x\) and we conclude by reflexivity of \(\rightarrow\). Otherwise \(t\) has shape \(f(\overline{t})\) and \(\overline{t} \rightarrow \overline{t}^{\sigma}\) by the IH and transitivity, so \(f(\overline{t}) \rightarrow f(\overline{t}^{\sigma})\). If the third clause applies we immediately conclude. Otherwise, \(f(\overline{t}^{\sigma}) = \ell^{r}\) and \(t^{\bullet} = r^{\sigma}\) for symbol \(f\), terms \(\overline{t}\), rule \(\ell \rightarrow r\) and substitution \(\sigma\), and we append \(\ell^{r} \rightarrow r^{\sigma}\);

(Rhs) We show the first by induction on \(t\). If \(t\) is a variable \(x\), then both sides are equal to \(\sigma(x)^{\bullet}\). Otherwise, \(t = f(\overline{t})\) for some symbol \(f\) and terms \(\overline{t}\), and \(f(\overline{t}^{\sigma}) \rightarrow (f(\overline{t}))^{\bullet}\) by the IH, hence \(f(\overline{t})(\sigma^{\bullet}) \rightarrow f((f(\overline{t}))^{\bullet})\). If the third clause applies to \(f(\overline{t})\) then we conclude, and otherwise we append a corresponding final root step to the reduction. For the second, note we have the stronger \(f(((f(\overline{t}))^{\bullet}))^{\bullet}) = f(f(\overline{t}))^{\bullet}\) in the induction step, so the second clause never applies as this is not an instance of a lhs by assumption on \(t\) and orthogonality;

(Z) We show for the full-development function \(\bullet\), that \(s \rightarrow t^{\bullet} \rightarrow s^{\bullet}\) for all steps \(t \rightarrow s\) by induction on \(t\). The case that \(t\) is a single variable being impossible, as variables cannot be rewritten due to the assumption that lhs of rules are not single variables, assume \(t\) has shape \(f(\overline{t})\) for symbol \(f\) and terms \(\overline{t}\) and distinguish cases on the clause of \(\bullet\).

Suppose the second clause applies, i.e. \(f(\overline{t}^{\sigma}) = \ell^{r}\) for some rule \(\ell \rightarrow r\) and \(t^{\bullet} = r^{r}\) for symbol \(f\), terms \(\overline{t}\), rule \(\ell \rightarrow r\) and substitution \(\tau\). Distinguish cases on the step \(t \rightarrow s\).

- If the step is a head step, then it must have shape \(t = \ell^{r} \rightarrow r^{\sigma} = s\) for the same rule \(\ell \rightarrow r\) and some substitution \(\sigma\) such that \(\sigma^{\bullet} = \tau\), by Rem. 21 and (Rhs) as \(t = f(\overline{t}) \rightarrow f(\overline{t}^{\bullet})\) by (Extensive). Then (Z) holds by \(r^{\sigma} \rightarrow r^{\tau} = (\ell^{r})^{\bullet} = r(\sigma^{\bullet}) = (r^{\sigma})^{\bullet}\) using (Extensive) for \(\sigma\) for the first reduction and (Rhs) for the second; and

- If the step is not a head step, then \(s = f(s)\) for some \(s\) equal to \(\ell^{\sigma}\) except for some \(i\) for which \(t_{i} \rightarrow s_{i}\), for which by the IH \(s_{i} \rightarrow t_{i}^{\bullet} \rightarrow s_{i}^{\bullet}\). From that, Rem. 21 and (Extensive) \(\ell^{r} = f(\overline{t}^{\bullet}) \rightarrow f(s^{\bullet}) = \ell^{r} \rightarrow r^{\sigma} = s^{\bullet}\) for some substitution \(\sigma\) with \(\tau \rightarrow \sigma\).

Using that for the second reduction, and the IH and (Extensive) for the first, (Z) holds by \(s = f(s)\) for \(f(\overline{t}^{\bullet}) = f(\ell)^{\bullet} \rightarrow r^{\sigma} = s^{\bullet}\).

Suppose the third clause applies, so \(t^{\bullet} = f(\overline{t})\). Then the step cannot be a head step (otherwise \(f(\overline{t}^{\bullet})\) would be a redex) and \(s = f(s)\) for some \(s\) equal to \(\ell^{\tau}\) except for some \(i\) for which \(t_{i} \rightarrow s_{i}\), for which by the IH \(s_{i} \rightarrow t_{i}^{\bullet} \rightarrow s_{i}^{\bullet}\). Then (Z) holds by using the IH and (Extensive) on \(\overline{t}\) for both reductions in \(f(s) \rightarrow f(t^{\bullet}) = f(\ell)^{\bullet} \rightarrow f(s^{\bullet})\), to which a further head step must be appended in case the second clause applies to \(s\) to yield \(s^{\bullet}\). □
In the proof of the lemma the condition \( f(\vec{t}) \) is a redex in the second clause of Def. 19 was never used. Indeed, dropping it preserves the proof. We dub the resulting function the full superdevelopment function as it relates to the full development function as Aczel’s proof of confluence [2, 26] relates to the Tait–Martin-Löf proof [4]; see [35] for a discussion. Full superdevelopments also contract all upward created [17] redexes.

Definition 23. Replacing redex by term in Def. 19 gives the full superdevelopment function.

Lemma 24. \( \rightarrow \) has the Z-property for the full superdevelopment function.

Example 25. Compared to Ex. 20 again \((I(Ix)) = x \) but now \((IIx) = x \) by also allowing to contract the upward created redex \( IX \). That CL has the Z-property is formalised in [9].

For simply typed CL we now already have seen 3 distinct functions witnessing the Z-property, in order of increasing(ly lax) upperbounds: full-development, full-superdevelopment, and normal form (Lem. 11 applies as simply typed CL is terminating).

3.3.2 Weakly orthogonal

We show weakly orthogonal TRSs [3, 38], having left-linear rules whose critical peaks \( s \leftarrow t \rightarrow r \) are trivial, i.e. \( s = r \), have the Z-property.

Example 26. The TRS with rules \( p(s(x)) \rightarrow x \) and \( s(p(x)) \rightarrow x \) is weakly orthogonal.

Definition 27. For a weakly orthogonal TRS, the maximal multistep map \( \bullet \) is inductively defined simultaneously with its maximal context \( \text{max} \) by

\[
\begin{align*}
x^\bullet &:= x \\
f(\vec{t})^\bullet &:= r^\sigma \\
\text{max}(x) &:= \Box \\
\text{max}(f(\vec{t})) &:= \Box \\
f(\hat{P}) &:= f(\text{max}(\vec{t})) \quad \text{otherwise}
\end{align*}
\]

where \( P \) asks \( f(\vec{t}) = \ell^\sigma \) for some substitution \( \sigma \), rule \( \ell \rightarrow r \) such that \( \ell \) is a prefix of \( f(\text{max}(\vec{t})) \).

Example 28. For the predecessor–successor TRS of Ex. 26 letting \( t := p(s(x)) \) and \( s := p(s(p(x))) \), we have \( \vec{t}^\bullet = x \) and \( \text{max}(t) = \Box \), respectively \( \vec{s}^\bullet = p(x) \) and \( \text{max}(s) = p(\Box) \).

The full development function being ambiguous\(^7\) for weakly orthogonal TRSs, is resolved by the maximal multistep map by adhering to an inside–out strategy. The intuition for \( \text{max}(t) \) is that it comprises the context of all maximal redexes selected for contraction by \( \bullet \), and the intuition for \( \bullet \) is that it tries to find any lhs that is contained in that context, i.e. does not have overlap with any of the already selected redexes in its arguments. As a consequence, in \( P \) the condition \( \ell \) is a prefix of \( f(\text{max}(\vec{t})) \) is always satisfied for TRSs that are orthogonal and for those the maximal multistep and full development functions coincide.

Lemma 29. \( \rightarrow \) has the Z-property for the maximal multistep function.

Proof. Since the Z-property is equivalent to the angle property, Lem. 8, this follows from the maximal multistep function having the angle property [38, Thm. 8.8.27], noting Def. 27 is a rephrasing of the notion going under the same name in the proof of that theorem. ▶

---

\(^7\) Different maximal sets of non-overlapping redexes may exist and result in different terms. E.g. the other redexes overlap the underlined one in \( p(s(p(s(x)))) \) hence the latter is maximal, but so are the other 2.
3.3.3 Associativity

From the above one might have the impression that the Z-property only holds for confluent TRSs that are orthogonal or closely associated to such. This is not the case.

Example 31. The term rewrite system for associativity (to the right) has as single rule:

\[ \alpha(\alpha(x, y), z) \rightarrow \alpha(x, \alpha(y, z)) \quad \text{xyz} \rightarrow x(yz) \]

written on the left in standard notation and applicatively (cf. Ex. 18) on the right.

As is well-known associativity is terminating and locally confluent as its one and only critical pair is joinable. Hence it has the Z-property by Cor. 17. Here we give a direct inductive definition of the normal form function, cf. Rem. 1, to show that one can proceed similarly to the (weakly) orthogonal case, and to prepare for the case of self-distributivity below.

Definition 32. We give an inductive definition of the normal form function ◦ depending on an auxiliary grafting function \( t[r] \) (we assume grafting binds stronger than the implicit \( \alpha \))

\[
\begin{align*}
x(r) &:= xr \\
(ts)(r) &:= ts[r] \\
(x)^* &:= x \\
(ts)^* &:= t^*(s^*)
\end{align*}
\]

The idea is that \( t[r] \) grafts the second argument \( r \) to the right tip of the first argument \( t \).

Example 33. \( (xy)^* = x^*(y^*) = xy \), so \( (xyz)^* = (xy)(z) = x(yz) \) and \( (xyzw)^* = x(yzw) \).

Note ◦ indeed only has normal forms in its image and these are preserved by grafting. The second example shows associativity can be viewed as performing an elementary case of grafting. How grafting and the normal form function interact with rewriting is captured by the following two lemmata, all of whose items are proven by induction on terms.\(^8\)

Lemma 34.

(Successionality) \( ts \rightarrow t[s] \), for all terms \( t,s \);  
(Compatible) \( t[s] \rightarrow t'[s'] \), if \( t \rightarrow t' \) and \( s \rightarrow s' \); and  
(Substitution) \( t[s][r] \rightarrow t[s')(r) \), for all terms \( t,s,r \).

Lemma 35.

(Extensive) \( t \rightarrow t^* \), for all terms \( t \);  
(Rhs) \( t^*(s^*r^*) \rightarrow (tsr)^* \), for all terms \( t,s,r \);  
(Z) \( \rightarrow \) has the Z-property for the normal form function ◦.

---

\(^8\) As shown there, this extends to infinitary rewriting, for non-collapsing TRSs.

\(^9\) See Appendix A to check that the proofs of the two lemmata are indeed by straightforward inductions.
Remark 36. Def. 32 effectively encodes a normalising strategy. A priori this entails neither termination of $\rightarrow$ nor uniqueness of the computed normal form.\textsuperscript{10} The latter only follows by the monotonicity part of the Z-property for $\bullet$. Turning things around, because $\bullet$ maps to normal forms, (Extensive) and monotonicity would have sufficed to establish the Z-property, as then $t \rightarrow s$ entails $s \rightarrow s^* = t^*$, but that would break the analogy with other proofs here.

3.3.4 Self-distributivity

Dehornoy’s proof that self-distributivity has the Z-property [6] fits in the above mould.

Example 37. The self-distributivity TRS has the (applicative) rule $xyz \rightarrow xz(yz)$.

Self-distributivity is non-terminating as its lhs can be embedded in its rhs, and is locally confluent as its one and only critical peak is joinable. Both its equational and rewrite theories are highly non-trivial; the book [6] is entirely devoted to them and still much more is to say.

Example 38. Self-distributivity has any ACI-operation (e.g., logical $\&$ or $\lor$) as model, as well as interpreting the binary operation as taking the middle between points in $\mathbb{R}^2$. The Substitution Lemma of the $\lambda$-calculus (cf. [32, Thm. 5]) yields an instance of self-distributivity. Self-distributivity is obtained by “forgetting” the $S$ in the CL rule for $S$, or alternatively (and giving more insight) by “enriching” the rhs of the associativity rule with another copy of $z$.

Definition 39. We give an inductive definition of the full distribution function $\bullet$ [6, Def. V.3.7] depending on the uniform distribution $t[s]$ of $s$ over $t$ [6, Def. V.3.4].

\begin{align*}
x[s] & := xs \\
(tr)[s] & := t[s][r][s] \\
x^* & := x \\
(ts)^* & := t^*[s^*]
\end{align*}

Uniform distribution grafts the 2nd argument uniformly to all leaves $t[s] = t[\ldots=x_1,x_2,\ldots]$. The following key lemmata, obtained by structuring [6, Lem. V.3.6,10–12] in the same way as was done for associativity above, are again proven by straightforward induction on terms.\textsuperscript{9}

Lemma 40. \begin{align*}
(\text{Sequentialisation}) & \quad ts \rightarrow t[s], \text{ for all terms } t,s; \\
(\text{Compatible}) & \quad t[s] \rightarrow t'[s'], \text{ if } t \rightarrow t' \text{ and } s \rightarrow s'; \text{ and} \\
(\text{Substitution}) & \quad t[s][r] \rightarrow t'[r][s[r]], \text{ for all terms } t,s,r.
\end{align*}

Lemma 41. \begin{align*}
(\text{Extensive}) & \quad t \rightarrow t^*, \text{ for all terms } t; \text{ and} \\
(\text{Z}) & \quad \rightarrow \text{ has the Z-property for the full distribution function } \bullet.
\end{align*}

3.4 The lambda-calculus

The $\lambda\beta$-calculus and the $\lambda\beta\eta$-calculus [4] being prime examples of orthogonal respectively weakly orthogonal higher-order term rewrite systems [20, 27], it is natural that the full development and full superdevelopment functions for orthogonal TRSs, and the maximal multistep map for weakly orthogonal TRSs should lift. They do. As the Z-property for the full development function is known [18]/[16] and for the full superdevelopment function was formalised [22, 9], we will be satisfied with giving the definitions and proof structure.

\textsuperscript{10}But in fact it can be shown to do so, by choosing appropriate weights in random descent [33].
Definition 42. The full development function $\bullet$ is inductively [37, p. 121] defined by:

$$
\begin{align*}
x^\bullet &:= x \\
(\lambda x.M)^\bullet &:= \lambda x.M^\bullet \\
(MN)^\bullet &:= M^\bullet[N^\prime/x^\prime] \quad \text{if } MN \text{ is a redex and } M^\bullet N^\prime = (\lambda x.M')N'
\end{align*}
$$

The full superdevelopment function is obtained by dropping the condition $MN$ is a redex from the third clause (or replacing it by $MN$ is a term; cf. Def. 19 and the text below Lem. 22).

Example 43. Taking $I := \lambda x.x$ in Ex. 20 gives full (super)developments as for CL.

Assuming $\alpha$-equivalence, congruence of $\beta$-reduction, the Substitution Lemma [4, Lem. 2.1.16], and compatibility of $\beta$-reduction with substitution [4, Sect. 3.1], and coherence of $\beta$-reduction with abstraction, we successively show:

Lemma 44.

(Extensive) $M \rightarrow M^\bullet$, for all $\lambda$-terms $M$;
(Rhs) $M^{\sigma^\bullet} \rightarrow (M^\bullet)^\sigma$ for $\lambda$-terms $M$, substitutions $\sigma$; and
(Z) $\rightarrow_\beta$ has the Z-property for the full (super)development function $\bullet$.

Remark 45. It would be interesting to see whether one could have a single formalised statement and proof for the Z-property for both full developments and full superdevelopments.

Remark 46. Our inside–out definition of the maximal multistep map for weakly orthogonal TRSs straightforwardly extends to all weakly orthogonal higher-order term rewrite systems, and the Z-property still holds (in [29] we established the angle property), which immediately yields the same for the $\lambda\beta\eta$-calculus. Although the outside–in construction on [37, p. 121, (F8∗)] does yield the Z-property for the $\lambda\beta\eta$-calculus, it fails to do so for weakly orthogonal higher-order term rewrite systems in general; monotonicity fails for the TRS in Rem. 30.

Remark 47. We do not know whether there is a generalisation of the full superdevelopment function to the $\lambda\beta\eta$-calculus. A problem is illustrated by the following example taken from [27, Rem. 3.4.24]. We have the co-initial full and non-full superdevelopments:

$$
(\lambda x.(\lambda y.y)x)I \rightarrow_\beta (\lambda x.L_x)z \rightarrow_\eta L_x \rightarrow_\beta z \quad (\lambda x.(\lambda y.y)x)I \rightarrow_\beta (\lambda y.yz)I
$$

but to reduce the target of the latter to that of the former requires two superdevelopments.

Example 48. The $\lambda$-calculus with explicit substitutions $\lambda\sigma$ [1] has the Z-property on closed terms. This is witnessed by the composition of first the function mapping a term to its $\rightarrow'$-normal form where $\rightarrow'$ denotes $\sigma$ reduction, and next the full development function $\bullet$ contracting all Beta-redexes (Beta on its own is orthogonal). The proof is given in Fig. 4, where black ordinary arrows denote Beta-reductions, blue open arrows $\rightarrow'$-reductions, $\tilde{t}$ the $\rightarrow'$-normal form of $t$, and $t^\bullet$ the result of subsequently applying the full-development function. For the result to hold, it suffices that

1. $\rightarrow'$ is confluent and terminating [38, Exercise 3.6.3(i)];
2. $\Rightarrow$ has the triangle property for $\bullet$; and
3. $\rightarrow'\ast$ single $\rightarrow'\ast$-steps commute with $\rightarrow'$-reduction [38, Exercise 3.6.3(iii)].

Example 49. We do not know whether Mints’ $\lambda$-calculus with restricted $\eta$-expansion (such that no $\beta$-redexes are created) has the Z-property. The restriction hampers monotonicity.

---

11 It coincides with the maximal multistep function since redex-clusters are chains [14,Defs. 4.31,4.47].
4 Syntax-free developments

We first show in Sects. 4.1–4.3 that several classical rewrite results that are known for the classical syntactic notion of development\(^\text{12}\) in term rewriting [38] and the \(\lambda\)-calculus [4] carry over to our syntax-free notion \(\bullet\rightarrow\) of \(\bullet\)-development (Def. 7) defined for a bullet function \(\bullet\) witnessing the Z-property. The diagrammatic proofs are obtained by pasting with Zs. Next, we investigate in Sect. 4.4 for the special case of orthogonal TRSs, under what conditions the syntactic and syntax-free notions of development coincide. Throughout we assume \(\rightarrow\) has the Z-property for \(\bullet\).

4.1 Hyper-Cofinality

We show \(\bullet\rightarrow\) is a best possible many-step strategy for \(\rightarrow\) in that it is hyper-cofinal [38, Sect. 9.1.1]; in order-theoretic terms: starting from object \(a\) and always eventually performing a \(\bullet\rightarrow\)-step eventually will yield a result greater than \(b\), for any \(b\) greater than \(a\). Observe first that \(\bullet\rightarrow\) is a many-step strategy since if \(a\rightarrow a^*\) then by Def. 4 \(a\) is not in \(\rightarrow\)-normal form, so there is some step \(a\rightarrow b\) from which we conclude to \(a\triangleright a^*\) by the Z-property.

\[\text{Theorem 50.} \quad \bullet\rightarrow\text{ is hyper-cofinal for } \rightarrow.\]

\[\text{Proof.} \quad \text{It suffices to show that, for a given step } a\rightarrow b \text{ and maximal [38, below Def. 1.1.13] reduction } \gamma \text{ of } \rightarrow, \bullet\rightarrow\text{-steps which always eventually contains a } \bullet\rightarrow\text{-step, there is another such reduction } \delta \text{ from } b \text{ eventually coinciding with it. By maximality, } \gamma \text{ either ends in a normal form } c, \text{ or by the assumption ("always eventually") decomposes into a } \rightarrow\text{-reduction } \gamma_1: a\rightarrow c, \text{ followed by } c\rightarrow c^* \text{ followed by another such reduction } \gamma_2 \text{ from } c^* \text{ (see Fig. 5). Induction on the length of } a\rightarrow c \text{ and monotonicity of } \bullet \text{ give a } d \text{ between } c \text{ and } c^* \text{ such that } \delta_1: b\rightarrow d. \text{ If } c \text{ is a normal form, } c = d \text{ and we set } \delta := \delta_1, \text{ else we compose } \delta \text{ from } \delta_1, d\rightarrow c^* \text{ and } \gamma_2.\]

\[\text{\textbf{Figure 5} Hyper-cofinality of } \rightarrow\text{ (left) and confluence of } \rightarrow\text{ (right), by tiling with Zs.}\]

\(^{12}\)Developments go all the way back to sequences of contractions on the parts in [5], for the \(\lambda I\)-calculus.
As a consequence [38, Sect. 9.1] $\bullet \rightarrow$ is a hyper-normalising strategy, i.e. if an object reduces to a normal form then always eventually performing a $\bullet \rightarrow$-step will reach it. For the $\lambda$-calculus $\bullet \rightarrow$ is (weak-)head-normalising, since (weak-)head-normal forms are closed under reduction; Normalisation of $\bullet \rightarrow$, i.e. of Gross–Knuth-reduction, was already noted in [18, Ex. 4.1].

4.2 Confluence

Lemma 51. $\rightarrow$ is confluent.

Proof. Confluence can be established in several ways. We present three.

- By tiling the plane with Zs as displayed in Fig. 5 (formally by the Strip Lemma and [38, Prop. 1.1.10]). In Fig. 5 we have highlighted the Zs for $a \rightarrow b$ and $a \rightarrow c$ in red and blue;
- Via Lem. 8, the angle property for $\leftrightarrow$ and [38, Prop. 1.1.11]; and
- Via Thm. 50, cofinality of $\rightarrow$ and [38, Thm. 1.2.3(iv)];\footnote{This generalises half of Staples’ confluence method [38, Exercise 1.3.9].}

$\bullet \rightarrow$ $\subseteq$ $\rightarrow$ $\bullet \rightarrow$ $\bullet \rightarrow$ $\subseteq$ $\rightarrow$ $\bullet \rightarrow$ $\bullet \rightarrow$ $\subseteq$ $\rightarrow$ $\bullet \rightarrow$

Since confluence is defined as the diamond property of the induced quasi-order, we have as a corollary that any rewrite system admitting the Z-property (Sect. 3.1) is confluent.

Remark 52. Choosing an appropriate bullet function (cf. Sect. 1) can lead to remarkably short proofs of confluence via the Z-property. To wit, the confluence proofs for positive braids (by full swaps), self-distributivity (by full distribution),\footnote{Confluence of self-distributivity is non-trivial. Currently no tool can prove it automatically; see problem 126 of http://cops.uibk.ac.at/results/?y=2020-full-run&c=TRS.} and for orthogonal TRSs and the $\lambda$-calculus (by full superdevelopments)\footnote{Full developments involve a useless test for being a redex (Def. 42).} are the shortest ones we know, in the same informal sense of “shortest” as was used by Takahashi on [37, p. 121] when she stated the proof of confluence of $\lambda\beta$ via the angle property was “perhaps the shortest”. However, the proof via the Z-property is (a bit) shorter [22].

Remark 53. Takahashi’s confluence proof method [37, Sect. 1] for the $\lambda$-calculus can be viewed as being based on the angle property for developments. Although the Z and angle properties are equivalent (Lem. 8), her method is slightly more involved, conceptually and technically, as it involves (inductively) defining both the bullet function and developments (called $*$ respectively parallel reduction in [37]). Our approach does away with the latter; our $\bullet$-developments are derived from $\bullet$ in a syntax-free way; beware though that developments and $\bullet$-developments in general differ, cf. Sect. 4.4.

4.3 Recurrence

[36, Proposition 1] characterises the recurrent terms in CL (see Ex. 18) in terms of Gross–Knuth reduction. We recast this in a syntax-free way for $\rightarrow$ having the Z-property.

Definition 54. An object $a$ is $\rightarrow$-recurrent if $a \rightarrow b$ entails $b \rightarrow a$ for all $b$. An object is recurrent if it is $\leftrightarrow$-recurrent.

Proposition 55. If $\bullet$ is extensive, then $a$ is recurrent iff $a^* \rightarrow a$.\footnote{This generalises half of Staples’ confluence method [38, Exercise 1.3.9].}
Proof. For the if-direction we show for all $n$, for all $b$, if $a \rightarrow^n b$ then $b \rightarrow a$, by induction on $n$. In the base case $a = b$ and we conclude by reflexivity of $\Rightarrow$. In the induction step, we have $a \rightarrow^n c \rightarrow b$ for some object $c$, so $c \rightarrow a$ by the IH for $a \rightarrow^n c$. We conclude by composing $b \rightarrow c^*$, which holds by the Z-property for $c \rightarrow b$, with $c^* \rightarrow a^*$, which holds by monotonicity of $\bullet$ for $c \rightarrow a$, and with $a^* \rightarrow a$, which holds by assumption, to $b \rightarrow a$ as desired.

For the only–if-direction, we have $a \rightarrow a^*$ by the assumption that $\bullet$ is extensive, hence $a^* \rightarrow a$ by the assumption that $a$ is recurrent, as desired. \hfill $\blacktriangleleft$

Remark 56. This result was used and formalised by Felgenhauer for a study of fixed-point combinators in CL [10]. E.g., although it is simple to see $\text{SII}(\text{SII})$ is recurrent, how to prove it in a simple way? By Proposition 55 it suffices to show that the result of a Gross–Knuth step reduces to it, i.e. that $I(\text{SII})(I(\text{SII})) \Rightarrow \text{SII}(\text{SII})$, which is simple to check.

### 4.4 Syntactic developments in orthogonal term rewriting

We investigate for orthogonal TRSs (cf. Sect. 3.3.1) the correspondence between the classical syntactic definition of a development and the syntax-free definition of $\bullet$-development (Def. 7) arising from taking as bullet function $\bullet$ the full development function that maps a term to the result of contracting all redex-patterns in it (Def. 19). This section is based on permutation equivalence via residual theory originating with [13], as presented in [38, Chs. 8 and 9]. We restrict to investigating the, non-trivial, correspondence for orthogonal TRSs hoping it can serve as a stepping stone for the same for more complex cases such as self-distributivity and the $\lambda$-calculus.

We first expand on the discrepancy between the syntactic and the syntax-free notions as observed in Ex. 9 (a non-terminating orthogonal TRS). Our first observation is that $\bullet$-developments are more encompassing than developments due to what are called syntactic accidents [17, p. 34], i.e. due to reductions yielding the same result despite not doing the same work, not being permutation equivalent. We show absence of syntactic accidents suffices.

Example 57. For the erasing TRS with rules $a \rightarrow b \rightarrow c$ and $f(x) \rightarrow d$, we have $f(a)^* := d$ and there is a $\bullet$-development from $f(a)$ to $f(c)$, but no such development. For the collapsing TRS with rules $g(x) \rightarrow h(x)$, $h(x) \rightarrow i(x)$ and $i(x) \rightarrow x$, we have $i(h(g(a)))^* := i(h(a))$ and there is a $\bullet$-development from $i(h(g(a)))$ to $i(h(i(a)))$, but no such development.

Proposition 58. For orthogonal, terminating, non-collapsing, and non-erasing TRSs, developments and $\bullet$-developments coincide.

Proof. We claim the assumptions guarantee the absence of syntactical accidents: if $\gamma, \delta$ are reductions from $t$ to $s$ then they are permutation equivalent $\gamma \simeq \zeta \delta$.\footnote{We employ the projection equivalence notation $\simeq$ from [38, Def. 8.7.21]. We freely employ results from that chapter, e.g. that permutation and projection equivalence coincide for orthogonal TRSs.} From the claim it follows that if $\gamma : t \rightarrow^{\bullet} t^*$ and $\delta : t \rightarrow s$ for some $\epsilon : s \rightarrow t^*$, then $\epsilon \simeq \zeta \delta \cdot \epsilon$. Therefore, decomposing $\delta$ as $\delta_1 \cdot \phi \cdot \delta_2$ for some step $\phi : t' \rightarrow s'$, we have $\gamma / \delta_1 : t' \rightarrow s$ and $\phi \subseteq \gamma / \delta_1$, which by non-erasiness entails that $\phi$ is among the redex-patterns in $\gamma / \delta_1$.\footnote{This fails for erasing systems. For instance, the step $f(a) \rightarrow f(s)$ is not a development of the step $f(a) \rightarrow c$ in the TRS with rules $a \rightarrow b$ and $f(x) \rightarrow c$.} Since this holds for each step, $\delta$ is a development of the set of all redex-patterns in $t$. The other implication follows from that every development from $t$ can be completed into a complete development to $t^*$.
It remains to prove the claim, which we prove by contradiction assuming \( \gamma \not\approx \delta \). By residual theory, the peak \( \gamma, \delta \) (where both have the same target, say \( u \), by accident) can be completed by a valley comprising \( \gamma' := \delta/\gamma \) and \( \delta' := \gamma/\delta \) such that \( \gamma \cdot \gamma' \simeq \delta \cdot \delta' \). At least one of \( \gamma', \delta' \) must be non-empty, as otherwise \( \gamma, \delta \) would be projection equivalent. But then the other must be non-empty as well, since otherwise we would have a reduction cycle on \( u \) contradicting the assumed termination. To see that \( \gamma' \not\approx \delta' \) note we may assume that \( \gamma, \delta \) are standard, where a reduction is standard [13] if for each step in it the position of the contracted redex-pattern is in the redex-pattern of the first step after and left–outer of it [38, Definition 8.5.40]. W.l.o.g. we may assume \( \gamma, \delta \) differ in their first steps and at least one of them contains a head-step, say \( \gamma \) contains head-step \( \phi \). Then \( \delta \) doesn’t, as otherwise their first steps would not differ by [15, Lemma 1]. We conclude \( \gamma/\delta \not\approx \delta/\gamma \) since the former contains a head-step as projection of a reduction \( \gamma \) containing a head-step over a reduction \( \delta \) containing none, and the latter contains no head-step as projection of \( \delta \) containing none over another reduction \( \gamma \) using the assumption that rules are non-collapsing. Applying the construction again, to the peak \( \gamma', \delta' \) (where both have the same target again by accident) yields a valley comprising \( \gamma'' := \delta'/\gamma' \) and \( \delta'' := \gamma'/\delta' \) such that \( \gamma' \cdot \gamma'' \simeq \delta' \cdot \delta'' \) but \( \gamma'' \not\approx \delta'' \). Repeating arbitrarily often yields an infinite reduction from \( t \), contradicting termination.

The three conditions in Prop. 58 are rather restrictive. We employ labelling [38, Sect. 8.4] to turn an arbitrary orthogonal term rewrite system into one satisfying them, and recover the result. We separate this into two phases, first turning a TRS into a non-erasing one by means of memorising the erased arguments,18 and next lifting to a TRS that is also terminating and non-collapsing by means of the Hyland–Wadsworth labelling [38, Sect. 8.4.4].

Definition 59. The TRS with memory \([T]\) of a TRS \( T \) has

- as signature the signature of \( T \) extended with a binary symbol \([, .]\);
- as rules \( g_j : \ell \rightarrow [r, x] \) for some \( T \)-rule \( g : \ell \rightarrow r \), where \( \ell \) is such that projecting all occurrences of \([, .] \) in it on their first argument yields \( \ell \), but these are not at the root, do not have a variable as first argument, and all have fresh variables (uniquely determined by their position) as second arguments. Here \( x \) is the list (unique for \( \ell \)) of all variables in \( \ell \) not in \( r \); \([t]\) denotes \( t \), and \([t, x]y\) denotes \([t, [x, y]]\).

Example 60. The TRS with memory for the rules \( f(a) \rightarrow b \) and \( f(x) \rightarrow b \), yields infinitely many rules \( f(a) \rightarrow b, f([a, x]) \rightarrow [b, x], f([a, y], x) \rightarrow [b, xy], \ldots \) for the first rule, and the single rule \( f(x) \rightarrow [b, x] \) for the second one.

Lemma 61. If \( T \) is orthogonal, then \([T]\) is orthogonal and non-erasing. The identity map induces a rewrite labelling [38, Def. 8.4.5(ii)] of \( T \) into \([T]\).

Example 62. Memorising overcomes erasingness. With memory \( f(a)^* := [d, b] \) for the first TRS in Ex. 57, so the \( \bullet \)-developments from \( f(a) \) are the initial prefixes of \( f(a) \rightarrow f(b) \rightarrow [d, b] \) and \( f(a) \rightarrow [d, a] \rightarrow [d, b] \). There is now no \( \bullet \)-development from \( f(a) \) to \( f(c) \).

To overcome also non-termination and (as a side-effect) collapsingness, we employ the Hyland–Wadsworth labelling [17, 4, 19, 38] \( T^\omega \) of a TRS \( T \). The idea of that labelling is to approximate arbitrary (possibly infinite) \( T \)-reductions with arbitrary precision, where precision is measured via the causal length of reductions. Technically, this is achieved by labelling edges19 in terms with their creation depth (a natural number) in such a way that any unlabelled reduction can be lifted to one having some bounded creation depth \( n \), and such that the corresponding subsystem \( T^n \) of \( T^\omega \) is terminating and confluent.

---

18 A technique going back to Nederpelt’s scars [24, p. 90].

19 To make sure that every redex-pattern contains at least one edge, we replace any function symbol \( f \) with a pair \( f' \cdot f \) with \( f' \) a fresh unary function symbol.
We have presented the Z-property and illustrated its flexibility, showing it applies to various rewrite systems to yield short proofs for classical results such as confluence and normalisation. Their proofs are based on a syntax-free version of the classical notion of development. We hope and expect more results can be factored in this way. We showed it coincides for orthogonalTRSs with the syntactic notion of development if syntactical accidents are absent (Prop. 58, Lem. 67) and hope that this invertibility result and its novel proof method extend to more complex systems, e.g. λ-calculus or self-distributivity.

5 Conclusion

We have presented the Z-property and illustrated its flexibility, showing it applies to various rewrite systems to yield short proofs for classical results such as confluence and normalisation. Their proofs are based on a syntax-free version of the classical notion of development. We hope and expect more results can be factored in this way. We showed it coincides for orthogonal TRSs with the syntactic notion of development if syntactical accidents are absent (Prop. 58, Lem. 67) and hope that this invertibility result and its novel proof method extend to more complex systems, e.g. λ-calculus or self-distributivity.
References


Proof of Lem. 34.

Proofs of second and third items of Lem. 8.

Assume $\rightarrow$ has the Z-property for bullet function $\bullet$. Define $\star$ to be $\bullet$ updated to map each object that is not the source of some step, to itself.

To see that $\star$ is extensive, we distinguish cases on whether $a$ is the source of some step or not. If it is, say $a \rightarrow b$, then $b \rightarrow a^\bullet \Rightarrow b^\bullet$ by the Z-property for $\bullet$. Hence $a \rightarrow a^\bullet = a^\star$ by composition and definition of $\star$. If it is not, then $a \rightarrow a^\bullet = a$ by reflexivity and definition of $\star$.

To see that $\rightarrow$ has the Z-property for $\star$, suppose $a \rightarrow b$. By the Z-property for $\bullet$ and by definition of $\star$, then $b \rightarrow a^\bullet = a^\star \rightarrow b^\star$. The result follows if, as we claim, $b^\bullet = b^\star$. That follows by noting that, by definition of $\star$, the only way in which $b^\bullet = b^\star$ could fail to hold, is if $b$ were not the source of some step. But then the above reduction collapses to $b = a^\bullet = a^\star = b^\star$ and we conclude since $b = b^\star$.

We only check the additional conditions on either side w.r.t. the first item.

For the only–if–direction, suppose $\rightarrow$ has the Z-property for an extensive $\bullet$. To show $a \rightarrow a^\bullet$, distinguish cases on whether there is some $\rightarrow$-step from $a$ or not. If there is, say $a \rightarrow b$ then by the Z-property, $a \rightarrow b^\bullet \Rightarrow a^\bullet$. If there is no $\rightarrow$-step from $a$, then extensivity of $\bullet$ entails $a = a^\bullet$. In either case, $a \rightarrow a^\bullet$ by reflexivity of $\rightarrow$, so $a \rightarrow a^\bullet$ by definition of $\rightarrow$.

Proof of Lem. 34.

(Sequentialisation) The proof is by induction on $t$. If $t$ is a variable, then $ts = t[s]$ and we conclude by reflexivity. Otherwise, $t$ has shape $t_1t_2$ and we conclude using the IH to $ts = t_1t_2s \Rightarrow t_1(t_2s) \Rightarrow t_1t_2[s] = t[s]$ from which the statement follows by transitivity.

(Compatible) We show the stronger fact that single steps in either $t$ or $s$ are preserved, by induction on $t$, which suffices by transitivity of $\rightarrow$. If $t$ is a variable $x$, then $s \rightarrow s'$ and $t[s] = xs \rightarrow xs' = t[s']$ by compatibility of reduction. If $t = t_1t_2$, we distinguish cases on where the step takes place:

- If the step takes place at the root of $t$, then $t = t_1t_2 \rightarrow t_1(t_2t_2) = t'$ and we conclude by unfolding the definition of right-substitution twice on both sides to $t[s] = t_1t_2t_2[s] \Rightarrow t_1(t_2t_2[s]) = t'[s]$;
- If the step takes place in $t_1$, then $t[s] = t_1t_2[s] \Rightarrow t_1t_2'[s] = t'[s]$ by compatibility of reduction;
- If the step takes place in $t_2$, then $t[s] = t_1t_2[s] \Rightarrow t_1t_2'[s] = t'[s]$ by the IH and compatibility of reduction;
- If the step takes place in $s$, then $t[s] = t_1t_2[s] \Rightarrow t_1t_2'[s] = t'[s]$ by the IH and compatibility of reduction.

(Substitution) The statement is shown by induction on $t$. If $t$ is a variable, say $x$ then $t[s][r] = xs[r] = t(s[r])$ by unfolding the definition of right-substitution. If $t$ has shape $t_1t_2$, then $t[s][r] = t_1t_2[s][r] = t_1t_2[s(r)] = t(s[r])$ by unfolding the definition of right-substitution and the IH.
Proof of Lem. 35.
(Extensive) By induction on t. If t is a variable, then \( t = t^* \) and we conclude by reflexivity of \( \rightarrow \). Otherwise \( t \) has shape \( t_1 t_2 \), and we conclude by (Sequentialisation), the IH twice, (Compatible), and definition to \( t_1 t_2 \rightarrow t_1 t_2 \equiv (t_1 t_2)^* \);
(Rhs) By (Sequentialization) twice and (Substitution) we conclude \( t^* (s^* r^*) \rightarrow t^* (s^* r^*) = t^* (s^* t^*) = (t s r)^* \);
(Z) As \( \bullet \) maps to normal forms, we show a strengthening of the Z-property, \( s \rightarrow t^* = s^* \), for all steps \( t \rightarrow s \), by induction and cases on \( t \).
If \( t \) is a variable, then the statement holds vacuously since the term then does not allow any step. Otherwise, \( t \) has shape \( t_1 t_2 \) and we distinguish cases on the position of the step.
- If the step takes place at the root, then \( t = (t_1 t_2) \rightarrow t_1 t_2 = s \), and we conclude using (extensive), (Rhs), the definition, and (Substitution) to \( t_1 t_2 \rightarrow t_1 t_2 = t_1 t_2 \equiv (t_1 t_2)^* \);
- If the step takes place in \( t_1 \), say \( t_1 \rightarrow s_1 \), then we conclude using the IH, (Extensive), (Sequentialisation) and definition to \( t_1 t_2 \rightarrow t_1 t_2 \equiv (t_1 t_2)^* = (s_1 t_2)^* \).
- If the step takes place in \( t_2 \) we proceed as in the previous item.
\( \blacktriangleleft \)

Proof of Lem. 40. Both items can be proven by induction on \( t \) or via the alternative definition of uniform distribution by means of substitution as given in the main text. We give samples of both:
(Sequentialisation) For variables \( xs = x[s] \), and for applications \( t_1 t_2 s \rightarrow t_1 s (t_2 s) \) we conclude using (Sequentialisation) first and then (Compatible) using the IH twice;
(Compatible) \( t[s] = t^* \) for the substitution \( \sigma \) mapping \( x \) to \( s \), and \( t'[s'] = t'^* \) for \( \sigma' \) mapping \( x \) to \( s' \). Hence if \( t \rightarrow t' \) and \( s \rightarrow s' \) then \( \sigma \rightarrow \sigma' \), hence \( t^* \rightarrow t'^* \) by compatibility of rewriting with substitution; and
(Substitution) For variables \( x[s][r] = (xs)[r] = xrs[r] \rightarrow x[r] [s[r]] \) by Sequentialisation twice, and for applications \( t_1 t_2 [s][r] = t_1 [s][r] t_2 [s][r] \rightarrow t_1 [r][s][r] t_2 [r][s][r] \equiv (t_1 t_2)[r][s][r] \) by the induction hypothesis twice.
\( \blacktriangleleft \)

Proof of Lem. 41. The items are proven by induction on \( t \).
(Extensive) For variables \( x = x^* \), and for applications \( ts \rightarrow t[s] \rightarrow t^*[s^*] = (ts)^* \) by (Sequentialisation) first and then (Compatible) using the IH twice;
(Z) We distinguish cases on whether the step is a head step or not.
- Suppose the step is a head step, so has shape \( tsr \rightarrow tr(sr) \). Then \( tr(sr) \rightarrow t[r][s][r] = (ts)[r] \rightarrow (ts)^*[r^*] = (tsr)^* \) by (Sequentialisation) and (Extensive), twice. Monotonicity of \( \bullet \) holds by \( (tsr)^* = t^*[s^*][r^*] \rightarrow t^*[r^*][s^*[r^*]] = (tr(sr))^* \) using (Substitution).
- If \( t_1 t_2 \rightarrow s_2 s_2 \) because \( t_1 \rightarrow s_i \) and \( t_{3-i} \) for some \( i \in \{1, 2\} \), then \( s_j \rightarrow t_j^* \sigma^* \) for \( j \in \{1, 2\} \), either by the IH, or (Extensive) and reflexivity. Using that, (Sequentialisation), and (Substitution) \( s_2 s_2 \rightarrow s_2 t_2 = (t_1 t_2)^* \leftrightarrow s_1^* t_2^* = (s_1 s_2)^* \).
\( \blacktriangleleft \)

Proof of Lem. 61. Orthogonality is preserved since brackets are only inserted between original function symbols, so overlapping \( \mathcal{T} \)-redex-patterns are mapped to overlapping \( \mathcal{T} \)-patterns by projecting brackets on their first arguments. That \( \mathcal{T} \) is non-erasing holds per construction.\(^{20}\)

The second part holds per construction of saturating left-hand sides of rules with memory.
\( \blacktriangleleft \)

\(^{20}\) if \( \mathcal{T} \) is orthogonal and right-linear, then \( \mathcal{T} \) is linear, so has random descent [30]: all reductions to a normal form have the same length.
Proof of Lem. 65. For the first item first note that its only–if-direction requires $n > 0$ as otherwise $T^n$ has no rules. Then, all (priming, labelling) operations for obtaining the rules of $T^n$ from those of $T$ are linear (only unary function symbols are added/removed) and redex-patterns overlapping in $T^n$ still do so after removing labels and collapsing $f’$–$f$-pairs to $f$. $T^n$ being a sub-system of $T^n$ the properties are preserved.

The second item holds per construction of the rules with both left- and right-hand sides being of shape $t'$ in which labels are inserted, for some $t$. Note that we also have the structural properties that reachable terms have at least one label between any two non-labels and removing all labels yields a term of shape $s'$ for some $s$.

Maranget [19] shows termination in the third item is a consequence of RPO, for the greater–than relation on labels, which is well-founded by the assumption that labels $< n$. Instead of basing ourselves on RPO, we can also give a direct inductive proof of termination in the style of van Daalen [17, 4]. In particular, we specialise the higher-order approach of [28] to first-order term rewriting. The proof is based on the so-called RHS lemma [28, Lemma 8] stating that a term rewrite system is terminating iff $r^\sigma$ is terminating for every rhs $r$ of a rule and terminating substitution $\sigma$. The only–if-direction of the RHS-lemma being trivial, to see the if-direction holds note that if there were a non-terminating term then there would be one of minimal size which then would have shape $f(\bar{t})$ with all $\bar{t}$ terminating by minimality. Hence an infinite reduction from it would have shape $f((\bar{t})) \rightarrow f(\bar{s}) = t^\sigma \rightarrow r^\sigma \rightarrow \ldots$ for some rule $\ell \rightarrow r$, substitution $\sigma$, and terms $\bar{s}$ such that $t_i \rightarrow s_i$ for all $i$. This is impossible as $r^\sigma$ is terminating by assumption since $\sigma$ is terminating as it assigns subterms of the $\bar{s}$ to variables and each $s_i$ is terminating as reduct of $t_i$.

To establish the assumption of the RHS lemma for $T^n$ we prove the more general claim that $(t^m)^\sigma$ is terminating for every $m \leq n$, term $t$ over (primed) symbols in $T$, and terminating substitution $\sigma$. This suffices as per construction of $T^n$ rhs of rules have this shape since labels in lhs are $< n$. The proof of the claim is by induction on the pair $(m, t)$ ordered by, in lexicographic order, the greater-than-or-equal order and the subterm order, and by distinguishing cases on the shape of $t$.

- If $t$ is a variable, then $(x^m)^\sigma := m(x^\sigma)$ and we conclude by the assumption that $\sigma$ is terminating, since the head symbol $m$ is not affected by any step per construction of $T^n$; labels occur in lhs only between (possibly primed) $T$-symbols.

- Otherwise $t$ has shape $f(\bar{t})$ for some (possibly primed) $T$-symbol. Since each $(t^m)^\sigma$ is terminating by the IH, which applies by a decrease in the second component of the pair, a hypothetical infinite reduction from $(t^m)^\sigma$ must then contain a head-step, i.e. have shape

$$m(f((t^m)^\sigma)) \rightarrow m(f(\bar{s})) = m(\ell^\sigma) \rightarrow m((r^k)^\tau) \rightarrow \ldots$$

for some $T^n$ rule of shape $\ell \rightarrow r^k$ with $k$ the maximum of the labels in $\ell$ plus one, substitution $\tau$ and $r$ a term over (possibly primed) $T$-symbols, and terms $\bar{s}$ such that $(t^m)^\sigma \rightarrow s_i$ for each $i$. This is impossible as $(r^k)^\tau$ is terminating by the IH, which applies by a decrease in the first component of the pair: $m < k$ because $(t^m)^\sigma \rightarrow s_i$ guarantees

---

21 Despite being intuitive and easy to prove the right-hand side lemma is informative: it would already fail for first-order TRSs if left-hand sides of rules were allowed to be single variables, consider the “rule” $x \rightarrow x$, and for higher-order TRSs it would fail if non-pattern-lhs were allowed [28].

22 Here we use that left-hand sides of term rewrite rules are not single variables.

23 To enable induction on terms; rhs of rules are not closed (as rhs!) under subterms in general.

24 Although terms of $T^n$ may contain labels $> n$, these need not be taken into consideration here. They have been “filtered-out” already by means of the RHS lemma so to speak, since labels $> n$ do not occur in the rules of $T^n$. 

---
that \( m \) is the head symbol of each \( s_i \), and per construction of \( T^\omega \) the lhs of any rule applicable to \( f(\vec{s}) \) contains the labels directly below \( f \); in fact \( f \) must be a (unary) primed symbol having a corresponding unprimed symbol (in \( T \)) below it. That \( \tau \) is terminating follows from that it assigns subterms of the \( s_i \) to variables, which are reducts of the \( (\iota^m_i)^\sigma \).

**Proof of Lem. 67.** We proceed as in the proof of Proposition 58, here stressing the similarity of structure and referring to that proof for details. We show that for all natural numbers \( n \), for all \( T^\omega_n \)-reductions \( \gamma, \delta : t \rightarrow s \) we have \( \gamma \simeq \delta \) by induction on \( t \) ordered by the union of \( \leftarrow \) and the sub-term relation, well-founded since \( T^\omega_n \) is terminating. This suffices since any pair of \( T^\omega \)-reductions is a pair of \( T^\omega_n \)-reductions (take \( n \) greater than all labels occurring in the redex-patterns contracted in \( \gamma, \delta \)), and \( T^\omega_n \)-projection equivalence entails \( T^\omega \)-projection equivalence.

Suppose \( \gamma, \delta \) were minimal such that \( \gamma \not\simeq \delta \). By residual theory, the peak \( \gamma, \delta \) can be completed by a valley comprising \( \gamma' := \delta/\gamma \) and \( \delta' := \gamma/\delta \) such that \( \gamma \cdot \gamma' \simeq \delta \cdot \delta' \). By assumption, at least one of \( \gamma', \delta' \) must be non-empty. We may assume that \( \gamma, \delta \) are standard, and by minimality that they don’t have the same first step (one may be empty), and at least one of them, say w.l.o.g. \( \gamma \), contains a head step. Since the system is orthogonal, [15, Lemma 1] yields then that \( \delta \) does not contain a head step. Hence \( \gamma/\delta \not\simeq \delta/\gamma \) since \( \gamma/\delta \) contains a head step as projection of a reduction \( \gamma \) containing a head step over a reduction \( \delta \) containing none, and \( \delta/\gamma \) contains no head step as projection of \( \delta \) containing none over another reduction \( \gamma \), using that \( T^\omega_n \)-rules are non-collapsing. Their sources being \( \rightarrow^+ \)-reachable from \( t \), \( \gamma/\delta, \delta/\gamma \) contradicts minimality of \( \gamma, \delta \).