Z

Syntax-Free Developments

Vincent van Oostrom ✉ 🄭

University of 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.

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 $\bullet$[3] mapping each object a to the chosen upper bound $a^\bullet$ of objects b such that $a \to b$, i.e. having a as single-step lower bound. Accordingly, we define the many-step rewrite *strategy* $\twoheadrightarrow\!\!\!\bullet$ to rewrite a into $a^\bullet$. For instance, taking as upper bound of a term t the term $t^\bullet$ obtained by a complete *development* of the *full* set of redexes in t, $\twoheadrightarrow\!\!\!\bullet$ is known as the *Gross–Knuth/full substitution* strategy in the λ-calculus/term rewriting [4, 38]. Based on $\bullet$, the classical notion of a

---

[1] [25] employs the reverse order, so speaks of existence of lower bounds.
[2] Newman leaves studing 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 $\twoheadrightarrow\!\!\!\bullet$ is bullet-fast; cf. Sect. 4.1.

development [5, 4, 38] can be given a *syntax-free* definition as $a \multimap b$ if $a \twoheadrightarrow b \twoheadrightarrow a^\bullet$; that is, *a develops* to b if b is *between* a and $a^\bullet$; with our notations suggesting that $\multimap$ is a development that is not as *full* as $\rightarrowtail$ 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 $\rightarrowtail$ and $\multimap$. 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, $\rightarrowtail$ can be seen as a *medium* step semantics; although $\rightarrowtail$-steps need not directly yield a normal form, they are monotonic. This may be suitable in a setting where for a step $a \to b$, 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* src, 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 $a, b, \dots$ range over objects, and $\phi, \psi, \dots$ 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.
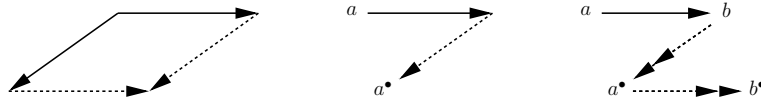
▭ **Figure 1** Diamond, angle, and Z-property for bullet function $\bullet$; named after the diagram shapes

▶ **Definition 4.** *A rewrite system $\to$ 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 $b \twoheadrightarrow a^\bullet$ for every a and step $a \to b$; and*
- ▬ Z *property if there is a map $\bullet$ such that $b \twoheadrightarrow a^\bullet \twoheadrightarrow b^\bullet$ for every a and step $a \to b$.*

where $\twoheadrightarrow$ denotes reduction, *finite (possibly empty) composition of steps. A map $\bullet$ is* extensive *if $a \twoheadrightarrow a^\bullet$ for all a, and induces a rewrite system $\rightarrowtail$ having the same objects as $\to$ and steps $a \rightarrowtail a^\bullet$ for all a not in $\to$-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 *tri*angle 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 diverse venues, e.g. [7, 31]. It has been introduced both before, for the λ-calculus, in [18, Ex. 4.1] and after in [16]. In the meantime it has been formalised [9] and applied, e.g. [23, 11].

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 $a \twoheadrightarrow b$ then $a^\bullet \twoheadrightarrow b^\bullet$ (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 $\to$ and map $\bullet$ on its objects, the $\bullet$-development rewrite system $\multimap\!\!\to$ has the objects of $\to$ and a step $a \multimap\!\!\to b$ for each pair of $\to$-reductions $a \twoheadrightarrow b \twoheadrightarrow a^\bullet$.*

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

▶ **Lemma 8.** *Let $\to$ be a rewrite system.*
- *$\to$ has the Z-property iff some $\to'$ such that $\to \subseteq \to' \subseteq \twoheadrightarrow$ has the angle property;[4]*
- *if $\to$ has the Z-property for $\bullet$, then it has the Z-property for some extensive $\star$; and*
- *$\to$ has the Z-property for an extensive $\bullet$ iff some rewrite system $\to'$ such that $\to \subseteq \to' \subseteq \twoheadrightarrow$ has the angle property and $a \to' a^\bullet$ 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 $\to'$ has the angle property, $\to \subseteq \to' \subseteq \twoheadrightarrow$, and suppose $a \to b$. Then by $\to \subseteq \to'$ and the angle property for $a \to' b$ we have $b \to' a^\bullet$, hence $a^\bullet \to' b^\bullet$ by applying the angle property again. Two angles make a Z; using $\to' \subseteq \twoheadrightarrow$ twice, we conclude to $b \twoheadrightarrow a^\bullet$ and $a^\bullet \twoheadrightarrow b^\bullet$.
  For the only–if-direction, assume $\to$ has the Z-property. Consider the $\bullet$-development rewrite system $\multimap\!\!\to$. To show $\multimap\!\!\to$ has the angle property, suppose $a \multimap\!\!\to b$. By definition $a \twoheadrightarrow b \twoheadrightarrow a^\bullet$. Combining $b \twoheadrightarrow a^\bullet$ with $a^\bullet \twoheadrightarrow b^\bullet$, which follows from $a \twoheadrightarrow b$ by monotonicity of $\bullet$, yields $b \multimap\!\!\to a^\bullet$ by definition of $\multimap\!\!\to$, showing the angle property. That the first inclusion in $\to \subseteq \multimap\!\!\to \subseteq \twoheadrightarrow$ holds follows from that $a \to b$ entails $b \twoheadrightarrow a^\bullet$ by the Z-property hence by definition $a \multimap\!\!\to b$, and that the second inclusion holds from that $a \multimap\!\!\to b$ unfolds to $a \twoheadrightarrow b \twoheadrightarrow a^\bullet$.
- 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 $\to' \subseteq \twoheadrightarrow$ 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 $a_i \to a_{i+1 \bmod 4}$ has the Z-property for the function
• mapping a_i to $a_{i+1 \bmod 4}$ because → is deterministic. Classically there are only two
developments from a_0 namely to itself, the empty development, and to a_1. However, because
→ is cyclic there are more •-developments, e.g. $a_0 \multimap\!\!\!\to a_2$ (since $a_0 \twoheadrightarrow a_2 \twoheadrightarrow a_1 = a_0^\bullet$).

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 → 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 $a \to b$ and $a \to c$, then $a = b$) systems by
mapping to *the next object* was exemplified in Ex. 9.

▶ **Lemma 10.** *If → is deterministic, then it has the Z-property.*

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

▶ **Lemma 11.** *If → is terminating, then → has the Z-property iff → is locally confluent.*

Proof. Suppose → is locally confluent and terminating. Let • be the *normal form* function
mapping each object to its →-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 $a \to b$ then $b \twoheadrightarrow a^\bullet = b^\bullet$. Vice versa, if → has the
Z-property for • then $a^\bullet$ is a common reduct to all b such that $a \to b$.   ◀

Ex. 6 shows there are confluent rewrite systems → that do not *have* the Z-property but
admit it in that there is a rewrite system →′ presenting the same quasi-order, i.e. $\twoheadrightarrow = \twoheadrightarrow'$,
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 $\to \subseteq \to' \subseteq \twoheadrightarrow$ entails →
and →′ present the same quasi-order. But there are confluent rewrite systems not admitting
either.

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

▶ Remark 13. Bullet functions for the Z-property may be *incomparable* (comparing their
images bulletwise by $\twoheadrightarrow$), 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].
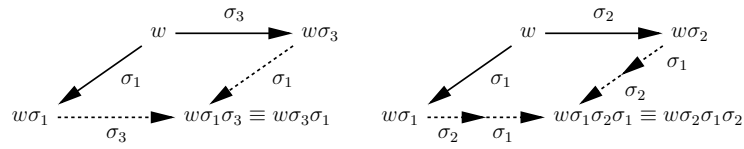
Figure 2 Local confluence diagrams for positive braids
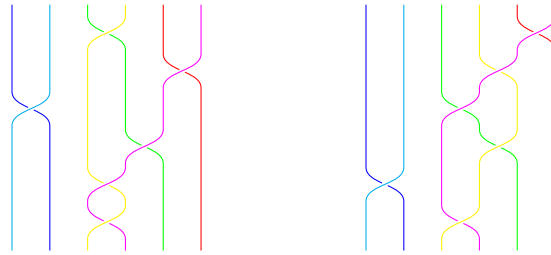
Figure 3 Isotopic braids, $\sigma_3\sigma_5\sigma_1\sigma_4\sigma_3\sigma_3 \equiv \sigma_5\sigma_4\sigma_3\sigma_4\sigma_1\sigma_3$, deformable into one another

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 ℓ strands has:*

▬ *as objects braids, words over the* Artin *generators σ_i for $1 \le i < \ell$, modulo*

$$\sigma_i\sigma_j\sigma_i = \sigma_j\sigma_i\sigma_j \qquad\qquad\qquad \textit{if } |i-j| = 1 \qquad\qquad (1)$$
$$\sigma_i\sigma_j = \sigma_j\sigma_i \qquad\qquad\qquad\quad \textit{if } |i-j| > 1 \qquad\qquad (2)$$

▬ *steps $w \to w\sigma_i$ for any braid w and $1 \le i < \ell$.*

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 σ_i, σ_j either is *too far apart* (2) like σ_1 and σ_3 on the left, or *too close together* (1) like σ_1 and σ_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 $\bullet$ suffixing the Garside word is formally defined by $w^\bullet := wG_\ell$, where, starting crossing from the left, the Garside word may be inductively defined by $G_0 := \varepsilon$ and if $n > 0$, then $G_n := G_{n\dot-1}\sigma_{\langle n,1\rangle}$ with $\sigma_{\langle i,j\rangle} := \sigma_{i\dot-1}\ldots\sigma_j$ crossing the ith strand over $i \dot- j$ strands to its left. The key property of G_ℓ 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 $1 \le i < n$ there exists a braid G_n^i such that (cf. Ex. 16)

$$\sigma_i G_n^i \equiv G_n \equiv G_n^i \sigma_{n\dot-i} \qquad\qquad\qquad\qquad\qquad\qquad (3)$$

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

It remains to prove the claim (a well-known fact). The intuition for G_n^i is that it is the *residual* of G_n after σ_i, i.e. what remains to be done of a *full* swap after swapping i. Formally, it may be inductively defined by $G_n^{n \dot- 1} := G_{n \dot- 1}\sigma_{\langle n,2\rangle}$ and $G_n^i := G_{n \dot- 1}^i\sigma_{\langle n,1\rangle}$ otherwise. Accordingly, we show (3) by induction on n, with trivial base case, and cases on whether or not $i = n \dot- 1$:

$$
\begin{array}{rclcrcl}
\sigma_i G_n^i & = & \sigma_i G_{n \dot- 1}^i\sigma_{\langle n,1\rangle} & \qquad & \sigma_{n \dot- 1} G_n^{n \dot- 1} & = & \sigma_{n \dot- 1} G_{n \dot- 2}\sigma_{\langle n \dot- 1,1\rangle}\sigma_{\langle n,2\rangle} \\
& \equiv_{IH} & G_{n \dot- 1}\sigma_{\langle n,1\rangle} & & & \equiv_{(i)} & G_{n \dot- 2}\sigma_{\langle n,1\rangle}\sigma_{\langle n,2\rangle} \\
& \equiv_{IH} & G_{n \dot- 1}^i\sigma_{n \dot- 1 \dot- i}\sigma_{\langle n,1\rangle} & & & \equiv_{(iii)} & G_{n \dot- 2}\sigma_{\langle n \dot- 1,1\rangle}\sigma_{\langle n,1\rangle} \\
& \equiv_{(ii)} & G_{n \dot- 1}^i\sigma_{\langle n,1\rangle}\sigma_{n \dot- i} & & & = & G_{n \dot- 1}\sigma_{\langle n,1\rangle} \\
& = & G_n^i\sigma_{n \dot- i} & & & = & G_n^{n \dot- 1}\sigma_1
\end{array}
$$

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

$$
\begin{array}{rcl}
\sigma_k \sigma_{\langle i,j\rangle} & \equiv_{(2)} & \sigma_{\langle i,k+2\rangle}\sigma_k\sigma_{k+1}\sigma_k\sigma_{\langle k,j\rangle} \\
& \equiv_{(1)} & \sigma_{\langle i,k+2\rangle}\sigma_{k+1}\sigma_k\sigma_{k+1}\sigma_{\langle k,j\rangle} \quad \equiv_{(2)} \quad \sigma_{\langle i,j\rangle}\sigma_{k+1}
\end{array}
$$

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

▶ **Example 16.** To see that (3) holds for $i := 2$ and $n := 4$, we first compute $G_4^2 := \sigma_1\sigma_2\sigma_3\sigma_2\sigma_1$ and $G_4 := \sigma_1\sigma_2\sigma_1\sigma_3\sigma_2\sigma_1$, and then verify $\underline{\sigma_2\sigma_1\sigma_2}\sigma_3\sigma_2\sigma_1 \equiv_{(1)} \sigma_1\sigma_2\underline{\sigma_1\sigma_3}\sigma_2\sigma_1 \equiv_{(2)} \sigma_1\sigma_2\sigma_3\underline{\sigma_1\sigma_2\sigma_1} \equiv_{(1)} \sigma_1\sigma_2\sigma_3\sigma_2\sigma_1\sigma_2$.

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 $\mathcal{T}$ is a TRS and $\to_\mathcal{T}$ or simply $\to$ 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):

$$
\begin{array}{rclcrcl}
@(I,x) & \to & x & \qquad\qquad & Ix & \to & x \\
@(@(K,x),y) & \to & x & & Kxy & \to & x \\
@(@(@(S,x),y),z) & \to & @(@(x,z),@(y,z)) & & Sxyz & \to & xz(yz)
\end{array}
$$

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].

▶ **Definition 19.** *For an orthogonal TRS,* full development • *is inductively defined by*

$$x^\bullet := x$$
$$f(\vec{t})^\bullet := r^\sigma \quad \text{if } f(\vec{t}) \text{ is a redex and } f(\vec{t}^{\,\bullet}) = \ell^\sigma \text{ for some rule } \ell \to r \text{ and substitution } \sigma$$
$$\quad\quad := f(\vec{t}^{\,\bullet}) \quad \text{otherwise}$$

▶ **Example 20.** In CL $(I(Ix))^\bullet = x$ and $(IIx)^\bullet = Ix$ contracting II but not the *created Ix*.

▶ Remark 21. By orthogonality, if for some *redex t* there is a reduction without head-steps $t \twoheadrightarrow \ell^\tau$ for lhs of a rule ℓ and substitution τ, then $t = \ell^\sigma$ for some substitution σ such that $\sigma \twoheadrightarrow \tau$. Vice versa, if we have such reduction $\ell^\tau \twoheadrightarrow t$ for some *term t*, then $t = \ell^\sigma$ and $\tau \twoheadrightarrow \sigma$.

▶ **Lemma 22.** *(Extensive) $t \twoheadrightarrow t^\bullet$ for all terms t;*
(Rhs) $t^{(\sigma^\bullet)} \twoheadrightarrow (t^\sigma)^\bullet$ for terms t, substitutions σ; $t^{(\sigma^\bullet)} = (t^\sigma)^\bullet$ if t is a proper subterm of a lhs;
(Z) $\to$ has the Z-property for the full development function.

Proof. (Extensive) By induction on t. If t is a variable x, then $t^\bullet = x$ and we conclude by reflexivity of $\twoheadrightarrow$. Otherwise t has shape $f(\vec{t})$ and $\vec{t} \twoheadrightarrow \vec{t}^{\,\bullet}$ by the IH and transitivity, so $f(\vec{t}) \twoheadrightarrow f(\vec{t}^{\,\bullet})$. If the third clause applies we immediately conclude. Otherwise, $f(\vec{t}^{\,\bullet}) = \ell^\sigma$ and $t^\bullet = r^\sigma$ for symbol f, terms $\vec{t}$, rule $\ell \to r$ and substitution σ, and we append $\ell^\sigma \to r^\sigma$;
(Rhs) We show the first by induction on t. If t is some variable x, then both sides are equal to $\sigma(x)^\bullet$. Otherwise, $t = f(\vec{t})$ for some symbol f and terms $\vec{t}$, and $\vec{t}^{(\sigma^\bullet)} \twoheadrightarrow (\vec{t}^\sigma)^\bullet$ by the IH, hence $f(\vec{t})^{(\sigma^\bullet)} \twoheadrightarrow f((\vec{t}^\sigma)^\bullet)$. If the third clause applies to $f(\vec{t}^\sigma)$ then we conclude, and otherwise we append a corresponding final root step to the reduction. For the second, note we have the stronger $f((\vec{t}^\sigma)^\bullet) = f(\vec{t})^{(\sigma^\bullet)}$ 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 •, that $s \twoheadrightarrow t^\bullet \twoheadrightarrow s^\bullet$ for all steps $t \to s$ by induction on t. The case that t is a single variable being impossible, as variables cannot be rewritten due to the assumption that lhss of rules are not single variables, assume t has shape $f(\vec{t})$ for symbol f and terms $\vec{t}$ and distinguish cases on the clause of •.
Suppose the second clause applies, i.e. $f(\vec{t}^{\,\bullet}) = \ell^\tau$ for some rule $\ell \to r$ and $t^\bullet = r^\tau$ for symbol f, terms $\vec{t}$, rule $\ell \to r$ and substitution τ. Distinguish cases on the step $t \to s$.
  - If the step is a head step, then it must have shape $t = \ell^\sigma \to r^\sigma = s$ for the *same* rule $\ell \to r$ and some substitution σ such that $\sigma^\bullet = \tau$, by Rem. 21 and (Rhs) as $t = f(\vec{t}) \twoheadrightarrow f(\vec{t}^{\,\bullet})$ by (Extensive). Then (Z) holds by $r^\sigma \twoheadrightarrow r^\tau = (\ell^\sigma)^\bullet = r^{(\sigma^\bullet)} \twoheadrightarrow (r^\sigma)^\bullet$ using (Extensive) for σ for the first reduction and (Rhs) for the second; and
  - If the step is not a head step, then $s = f(\vec{s})$ for some $\vec{s}$ equal to $\vec{t}$ except for some i for which $t_i \to s_i$, for which by the IH $s_i \twoheadrightarrow t_i^\bullet \twoheadrightarrow s_i^\bullet$. From that, Rem. 21 and (Extensive) $\ell^\tau = f(\vec{t}^{\,\bullet}) \twoheadrightarrow f(\vec{s}^{\,\bullet}) = \ell^\sigma \to r^\sigma = s^\bullet$ for some substitution σ with $\tau \twoheadrightarrow \sigma$. Using that for the second reduction, and the IH and (Extensive) for the first, (Z) holds by $f(\vec{s}) \twoheadrightarrow f(\vec{t}^{\,\bullet}) = \ell^\tau \to r^\tau = f(\vec{t})^\bullet \twoheadrightarrow r^\sigma = s^\bullet = f(\vec{s})^\bullet$.
Suppose the third clause applies, so $t^\bullet = f(\vec{t}^{\,\bullet})$. Then the step cannot be a head step (otherwise $f(\vec{t}^{\,\bullet})$ would be a redex) and $s = f(\vec{s})$ for some $\vec{s}$ equal to $\vec{t}$ except for some i for which $t_i \to s_i$, for which by the IH $s_i \twoheadrightarrow t_i^\bullet \twoheadrightarrow s_i^\bullet$. Then (Z) holds by using the IH and (Extensive) on $\vec{t}$ for both reductions in $f(\vec{s}) \twoheadrightarrow f(\vec{t}^{\,\bullet}) = f(\vec{t})^\bullet \twoheadrightarrow f(\vec{s}^{\,\bullet})$, 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

*super*development 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* super*development function.*

▶ **Lemma 24.** $\to$ *has the Z-property for the full superdevelopment function.*

▶ **Example 25.** Compared to Ex. 20 again $(I(Ix))^\bullet = x$ but now $(IIx)^\bullet = x$ 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 $s \leftarrow t \to r$ are *trivial*, i.e. $s = r$, have the Z-property.

▶ **Example 26.** The TRS with rules $p(s(x)) \to x$ and $s(p(x)) \to x$ is weakly orthogonal.

▶ **Definition 27.** *For a weakly orthogonal TRS, the* maximal multistep *map* $\bullet$ *is inductively defined simultaneously with its* maximal context max *by*

$$
\begin{array}{rclcrcll}
x^\bullet & := & x & & \max(x) & := & \square & \\
f(\vec{t})^\bullet & := & r^\sigma & & \max(f(\vec{t})) & := & \square & \textit{if } P \\
& := & f(\vec{t}^\bullet) & & & := & f(\max(\vec{t})) & \textit{otherwise}
\end{array}
$$

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

▶ **Example 28.** For the predecessor–successor TRS of Ex. 26 letting $t := p(s(x))$ and $s := p(s(p(x)))$, we have $t^\bullet = x$ and $\max(t) = \square$, respectively $s^\bullet = p(x)$ and $\max(s) = p(\square)$.

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 $\max(t)$ 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 ℓ *is a prefix of* $f(\max(\vec{t}))$ is always satisfied for TRSs that are orthogonal and for those the maximal multistep and full development functions coincide.

▶ **Lemma 29.** $\to$ *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 $p(\underline{s(p(s(x))))}$ hence the latter is maximal, but so are the other 2.

▶ **Remark 30.** Proceeding *outside–in* instead of *inside–out*, in a naïve way cannot work. It does not yield a bullet function having the Z-property as exemplified by the TRS with rules $c(x) \to x$, $f(f(x)) \to f(x)$ and $g(f(f(f(x)))) \to g(f(f(x)))$. We have $t \to s$ for $t := g(f(f(c(f(f(x))))))$ and $s := g(f(f(f(f(x)))))$ by contracting the c-redex, but the Z-property (monotonicity) fails for a naïve outside–in bullet function $\star$, as we do *not* have $t^\star = g(f(f(x))) \twoheadrightarrow g(f(f(f(x)))) = s^\star$. This can be overcome [8, Lem. 7.10][8], even effectively so [8, Cor. 7.27], by discarding *Takahashi configurations* [38, Prop. 9.3.5], [14, Rem. 4.38].

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:

$$@(@(x, y), z) \quad \to \quad @(x, @(y, z)) \qquad xyz \quad \to \quad x(yz)$$

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* $\bullet$ *depending on an auxiliary* grafting *function* $t\langle r]$ *(we assume grafting binds stronger than the implicit* @*)*

$$
\begin{array}{rclcrcl}
x\langle r] & := & xr & \qquad & x^\bullet & := & x \\
(ts)\langle r] & := & ts\langle r] & & (ts)^\bullet & := & t^\bullet\langle s^\bullet]
\end{array}
$$

The idea is that $t\langle r]$ grafts the second argument r to the right tip of the first argument t.

▶ **Example 33.** $(xy)^\bullet = x^\bullet\langle y^\bullet] = xy$, so $(xyz)^\bullet = (xy)\langle z] = x(yz)$ and $(xyzw)^\bullet = x(y(zw))$.

Note $\bullet$ 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.[9]

▶ **Lemma 34.** *(Sequentialisation)* $ts \twoheadrightarrow t\langle s]$, *for all terms* t, s;
(Compatible) $t\langle s] \twoheadrightarrow t'\langle s']$, *if* $t \twoheadrightarrow t'$ *and* $s \twoheadrightarrow s'$; *and*
(Substitution) $t\langle s]\langle r] = t\langle s\langle r]]$, *for all terms* t, s, r.

▶ **Lemma 35.** *(Extensive)* $t \twoheadrightarrow t^\bullet$, *for all terms* t;
(Rhs) $t^\bullet(s^\bullet r^\bullet) \twoheadrightarrow (tsr)^\bullet$, *for all terms* t, s, r;
(Z) $\to$ *has the Z-property for the normal form function* $\bullet$.

▶ **Remark 36.** Def. 32 effectively encodes a normalising *strategy*. *A priori* this entails neither termination of $\to$ nor uniqueness of the computed normal form.[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 $t \to s$ entails $s \twoheadrightarrow s^\bullet = t^\bullet$, but that would break the analogy with other proofs here.

---

[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.
[10] But in fact it can be shown to do so, by choosing appropriate weights in *random descent* [33].

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 $xyz \to xz(yz)$.

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 $\wedge$ or $\vee$) as model, as well as interpreting the binary operation as *taking the middle* between points in $\mathbb{R}^2$. The *Substitution Lemma* of the λ-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* • *[6, Def. V.3.7] depending on the* uniform *distribution* $t[s]$ *of* s *over* t *[6, Def. V.3.4].*

$$
\begin{array}{rclcrcl}
x[s] & := & xs & \quad & x^{\bullet} & := & x \\
(tr)[s] & := & t[s]r[s] & \quad & (ts)^{\bullet} & := & t^{\bullet}[s^{\bullet}]
\end{array}
$$

Uniform distribution grafts the 2nd argument uniformly to all leafs $t[s] = t^{[x_1, x_2, \ldots := x_1 s, x_2 s, \ldots]}$. 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.[9]

▶ **Lemma 40.**   *(Sequentialisation)* $ts \twoheadrightarrow t[s]$, *for all terms* t, s;
(Compatible) $t[s] \twoheadrightarrow t'[s']$, *if* $t \twoheadrightarrow t'$ *and* $s \twoheadrightarrow s'$; *and*
(Substitution) $t[s][r] \twoheadrightarrow t[r][s[r]]$, *for all terms* t, s, r.

▶ **Lemma 41.**   *(Extensive)* $t \twoheadrightarrow t^{\bullet}$, *for all terms* t; *and*
(Z) $\to$ *has the Z-property for the uniform distribution function* •.

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.

▶ **Definition 42.** *The* full development *function* • *is inductively [37, p. 121] defined by:*

$$
\begin{array}{rcll}
x^{\bullet} & := & x & \\
(\lambda x.M)^{\bullet} & := & \lambda x.M^{\bullet} & \\
(MN)^{\bullet} & := & M'^{[x := N']} & \textit{if } MN \textit{ is a redex and } M^{\bullet}N^{\bullet} = (\lambda x.M')N' \\
& := & M^{\bullet}N^{\bullet} & \textit{otherwise}
\end{array}
$$

The full super*development 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 α-equivalence, congruence of β-reduction, the Substitution Lemma [4, Lem. 2.1.16], and compatibility of β-reduction with substitution [4, Sect. 3.1], and coherence of β-reduction with abstraction, we successively show:

▶ **Lemma 44.** *(Extensive)* $M \twoheadrightarrow M^\bullet$, *for all λ-terms M;*
(Rhs) $M^{(\sigma^\bullet)} \twoheadrightarrow (M^\sigma)^\bullet$ for λ-terms M, substitutions σ; and
(Z) $\to_\beta$ 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,[11] 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:

$$(\lambda x.\underline{(\lambda y.yx)I})z \to_\beta (\underline{\lambda x.Ix})z \to_\eta \underline{Iz} \to_\beta z \qquad \underline{(\lambda x.(\lambda y.yx)I)z} \to_\beta (\lambda y.yz)I$$

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

▶ **Example 48.** The λ-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 $\to'$-normal form where $\to'$ denotes σ 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 $\to'$-reductions, $\bar{t}$ the
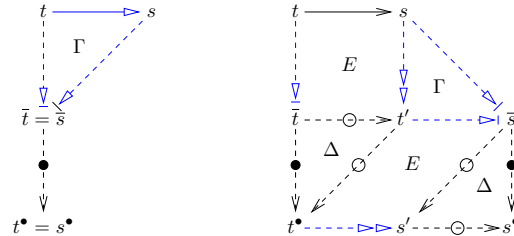
Figure 4 $\lambda\sigma$ has the Z-property

$\to'$-normal form of t, and $t^\bullet$ the result of subsequently applying the full-development function. For the result to hold, it suffices that

(Γ) $\to'$ is confluent and terminating [38, Exercise 3.6.3(i)];
(Δ) $\multimap\joinrel\to$ has the triangle property for $\bullet$; and
(E) single $\multimap\joinrel\to$-steps commute with $\to'$-reduction [38, Exercise 3.6.3(iii)].

▶ **Example 49.** We do not know whether Mints' λ-calculus with *restricted* η-expansion (such that no β-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[12] in term rewriting [38] and the λ-calculus [4] carry over to our syntax-free notion $\multimap\!\!\!\!\rightarrow$ 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 $\multimap\!\!\!\!\rightarrow$ 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 $\multimap\!\!\!\!\rightarrow$-step eventually will yield a result greater than b, for any b greater than a. Observe first that $\multimap\!\!\!\!\rightarrow$ *is* a many-step strategy since if $a \multimap\!\!\!\!\rightarrow a^\bullet$ then by Def. 4 a is not in $\rightarrow$-normal form, so there is some step $a \rightarrow b$ from which we conclude to $a \twoheadrightarrow a^\bullet$ by the Z-property.

▶ **Theorem 50.** $\multimap\!\!\!\!\rightarrow$ *is* hyper-cofinal *for* $\rightarrow$.

Proof. It suffices to show that, for a given step $a \rightarrow b$ and *maximal [38, below Def. 1.1.13] reduction γ of $\rightarrow,\multimap\!\!\!\!\rightarrow$-steps which always eventually contains a $\multimap\!\!\!\!\rightarrow$-step*, there is another such reduction δ from b eventually coinciding with it. By maximality, γ either ends in a
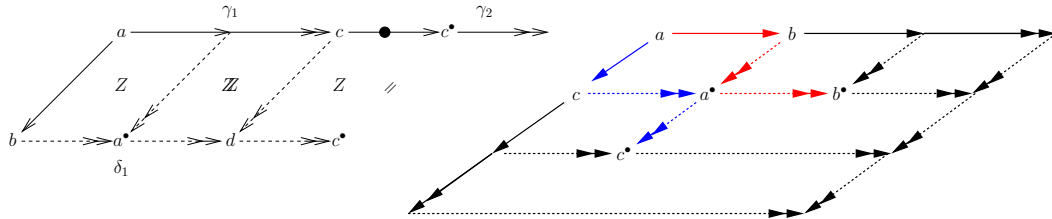
Figure 5 Hyper-cofinality of $\multimap\!\!\!\!\rightarrow$ (left) and confluence of $\rightarrow$ (right), by tiling with Zs

normal form c, or by the assumption ('always eventually') decomposes into a $\rightarrow$-reduction $\gamma_1 : a \twoheadrightarrow c$, followed by $c \multimap\!\!\!\!\rightarrow c^\bullet$ followed by another such reduction γ_2 from $c^\bullet$ (see Fig. 5). Induction on the length of $a \twoheadrightarrow c$ and monotonicity of $\bullet$ give a d between c and $c^\bullet$ such that $\delta_1 : b \twoheadrightarrow d$. If c is a normal form, $c = d$ and we set $\delta := \delta_1$, else we compose δ from δ_1, $d \twoheadrightarrow c^\bullet$ and γ_2.   ◀

As a consequence [38, Sect. 9.1] $\multimap\!\!\!\!\rightarrow$ is a *hyper-normalising* strategy, i.e. if an object reduces to a normal form then always eventually performing a $\multimap\!\!\!\!\rightarrow$-step will reach it. For the λ-calculus $\multimap\!\!\!\!\rightarrow$ is (weak-)head-normalising, since (weak-)head-normal forms are closed under reduction; Normalisation of $\multimap\!\!\!\!\rightarrow$, i.e. of Gross–Knuth-reduction, was already noted in [18, Ex. 4.1].

4.2   Confluence

▶ **Lemma 51.** $\rightarrow$ *is confluent.*

---

[12] Developments go all the way back to *sequences of contractions on the parts* in [5], for the λI-calculus.

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 high-lighted the Zs for $a \to b$ and $a \to c$ in red and blue;
- Via Lem. 8, the angle property for $\twoheadrightarrow$ and [38, Prop. 1.1.11]; and
- Via Thm. 50, cofinality of $\twoheadrightarrow\hspace{-0.6em}\bullet$ and [38, Thm. 1.2.3(iv)]:[13]

$$
\begin{aligned}
\twoheadleftarrow \cdot \twoheadrightarrow \ &\subseteq\ \twoheadrightarrow \cdot \twoheadleftarrow\hspace{-0.6em}\bullet\ \cdot\ \bullet\hspace{-0.6em}\twoheadrightarrow \cdot \twoheadleftarrow && \text{as } \to \subseteq \ \bullet\hspace{-0.6em}\twoheadrightarrow \cdot \twoheadleftarrow \text{ by cofinality of } \twoheadrightarrow\hspace{-0.6em}\bullet \\
&\subseteq\ \twoheadrightarrow \cdot \bullet\hspace{-0.6em}\twoheadrightarrow \cdot \twoheadleftarrow\hspace{-0.6em}\bullet\ \cdot \twoheadleftarrow && \text{as } \twoheadrightarrow\hspace{-0.6em}\bullet \text{ is deterministic hence confluent} \\
&\subseteq\ \twoheadrightarrow \cdot \twoheadrightarrow \cdot \twoheadleftarrow \cdot \twoheadleftarrow && \text{as } \twoheadrightarrow\hspace{-0.6em}\bullet \text{ is a many-step } \to\text{-strategy} \\
&\subseteq\ \twoheadrightarrow \cdot \twoheadleftarrow && \text{by transitivity of } \twoheadrightarrow \qquad\qquad\blacktriangleleft
\end{aligned}
$$

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),[14] and for orthogonal TRSs and the λ-calculus (by full superdevelopments)[15] 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 λ-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 $\to$ having the Z-property.

▶ **Definition 54.** *An object a is $\to$-recurrent if $a \to b$ entails $b \twoheadrightarrow a$ for all b. An object is recurrent if it is $\twoheadrightarrow$-recurrent.*

▶ **Proposition 55.** *If $\bullet$ is extensive, then a is recurrent iff $a^\bullet \twoheadrightarrow a$.*

Proof. For the if-direction we show for all n, for all b, if $a \to^n b$ then $b \twoheadrightarrow a$, by induction on n. In the base case $a = b$ and we conclude by reflexivity of $\twoheadrightarrow$. In the induction step, we have $a \to^n c \to b$ for some object c, so $c \twoheadrightarrow a$ by the IH for $a \to^n c$. We conclude by composing $b \twoheadrightarrow c^\bullet$, which holds by the Z-property for $c \to b$, with $c^\bullet \twoheadrightarrow a^\bullet$, which holds by monotonicity of $\bullet$ for $c \twoheadrightarrow a$, and with $a^\bullet \twoheadrightarrow a$, which holds by assumption, to $b \twoheadrightarrow a$ as desired.

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

---

[13] This generalises half of Staples' confluence method [38, Exercise 1.3.9].

[14] 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`.

[15] Full developments involve a useless test for being a redex (Def. 42).

▶ **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* $SII(SII)$ 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 $I(SII)(I(SII)) \twoheadrightarrow SII(SII)$, 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 •-development (Def. 7) arising from taking as bullet function • 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 λ-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 •-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 $a \to b \to c$ and $f(x) \to d$, we have $f(a)^\bullet := d$ and there is a •-development from $f(a)$ to $f(c)$, but no such development. For the *collapsing* TRS with rules $g(x) \to h(x)$, $h(x) \to i(x)$ and $i(x) \to x$, we have $i(h(g(a)))^\bullet := i(h(a))$ and there is a •-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 •-developments coincide.*

Proof. We claim the assumptions guarantee the absence of syntactical accidents: if γ, δ are reductions from t to s then they are permutation equivalent $\gamma \simeq \delta$.[16] From the claim it follows that if $\gamma : t \multimap\!\!\to t^\bullet$ and $\delta : t \twoheadrightarrow s$ for some $\epsilon : s \twoheadrightarrow t^\bullet$, then $\gamma \simeq \delta \cdot \epsilon$. Therefore, decomposing δ as $\delta_1 \cdot \phi \cdot \delta_2$ for some step $\phi : t' \to s'$, we have $\gamma/\delta_1 : t' \multimap\!\!\to s$ and $\phi \lesssim \gamma/\delta_1$, which by non-erasingness entails that ϕ is among the redex-patterns in γ/δ_1.[17] Since this holds for each step, δ 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 $t^\bullet$.

It remains to prove the claim, which we prove by contradiction assuming $\gamma \not\simeq \delta$. By residual theory, the peak γ, δ (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 γ', δ' must be non-empty, as otherwise γ, δ 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 $\gamma' \not\simeq \delta'$ note we may assume that γ, δ 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

---

[16] 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.

[17] This fails for erasing systems. For instance, the step $f(a) \to f(s)$ is not a development of the step $f(a) \to c$ in the TRS with rules $a \to b$ and $f(x) \to c$.

it [38, Definition 8.5.40]. W.l.o.g. we may assume γ, δ differ in their first steps and at least one of them contains a head-step, say γ contains head-step ϕ. Then δ doesn't, as otherwise their first steps would not differ by [15, Lemma 1]. We conclude $\gamma/\delta \not\simeq \delta/\gamma$ since the former contains a head-step as projection of a reduction γ containing a head-step over a reduction δ containing none, and the latter contains no head-step as projection of δ containing none over another reduction γ using the assumption that rules are non-collapsing. Applying the construction again, to the peak γ', δ' (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 $\gamma'' \not\simeq \delta''$. 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* $[\mathcal{T}]$ *of a TRS* $\mathcal{T}$ *has*

▬ *as signature the signature of* $\mathcal{T}$ *extended with a binary symbol* $[\,,\,]$;

▬ *as rules* $\varrho_{\hat{\ell}} : \hat{\ell} \to [r, \vec{x}]$ *for some* $\mathcal{T}$*-rule* $\varrho : \ell \to r$, *where* $\hat{\ell}$ *is such that projecting all occurrences of* $[\,,\,]$ *in it on their first argument yields* ℓ, *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* $\vec{x}$ *is the list (unique for* $\hat{\ell}$*) of all variables in* $\hat{\ell}$ *not in* r, $[t]$ *denotes* t, *and* $[t, x\vec{y}]$ *denotes* $[t, [x, \vec{y}]]$.

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

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

▶ **Example 62.** Memorising overcomes erasingness. With memory $f(a)^\bullet := [d, b]$ for the first TRS in Ex. 57, so the $\bullet$-developments from $f(a)$ are the initial prefixes of $f(a) \to f(b) \to [d, b]$ and $f(a) \to [d, a] \to [d, b]$. 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] $\mathcal{T}^\omega$ of a TRS $\mathcal{T}$. The idea of that labelling is to approximate arbitrary (possibly infinite) $\mathcal{T}$-reductions with arbitrary precision, where precision is measured via the causal length of reductions. Technically, this is achieved by labelling edges[19] 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 $\mathcal{T}^n$ of $\mathcal{T}^\omega$ is *terminating* and confluent.

▶ **Definition 63.** *The* Hyland–Wadsworth *(HW) labelling of a TRS* $\mathcal{T}$ *is the TRS* $\mathcal{T}^\omega$

▬ *having as signature all natural numbers (*labels*) and for every* f *of* $\mathcal{T}$ *both* f *and a fresh copy* f' *of it, with all symbols not in* $\mathcal{T}$ *having arity* 1;

---

[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 $f'-f$ with f' a fresh unary function symbol.

▬ *having as rules $\varrho_{\hat{\ell}} : \hat{\ell} \to (r')^n$ for every rule $\varrho : \ell \to r$, where $\hat{\ell}$ is such that between any two non-labels there is at least one label, n is the maximum value of all labels in $\hat{\ell}$ plus one, and removing all yields ℓ', where priming and natural-number-labelling are defined by:*

$$
\begin{array}{rclcrcl}
x' & := & x & & x^n & := & n(x) \\
f(\vec{t}\,)' & := & f'(f(\vec{t'})) & & g(\vec{s})^n & := & n(g(s^{\vec{n}}))
\end{array}
$$

$\mathcal{T}^n$ *is the restriction of $\mathcal{T}^\omega$ to rules whose lhss have labels $< n$.*

▶ **Example 64.** We illustrate the saturation process of the HW-labelling on a rule with a single-function-symbol left-hand side (cf. footnote 19), The Hyland–Wadsworth labelling of the TRS with rule $f(x) \to x$ has the infinitely many rules $f'(0(f(x))) \to 1(x)$, $f'(1(f(x))) \to 2(x)$, ..., $f'(0(0(f(x)))) \to 1(x)$, $f'(1(0(f(x)))) \to 2(x)$, $f'(0(1(f(x)))) \to 2(x)$, ... Note the original rule was collapsing, but its HW-labellings are not.

Hyland–Wadsworth labelling preserves orthogonality and is sound in that reductions can be lifted, however with ever increasing labels so bounding them yields termination.

▶ **Lemma 65.** ▬ *$\mathcal{T}^\omega$ and $\mathcal{T}^n$ are (left/right) linear and/or orthogonal iff $\mathcal{T}$ is;*

▬ *mapping every term t to $(t')^0$ gives a rewrite labelling of $\mathcal{T}$; and*

▬ *the restriction $\mathcal{T}^n$ of $\mathcal{T}^\omega$ to lhs with labels $< n$ is terminating [19].*

▶ **Example 66.** To see how the HW-labelling avoids syntactical accidents for collapsing rules consider the reduction $f(f(x)) \to f(x)$ for rule $f(x) \to x$. It lifts differently depending on which redex-pattern is contracted:

$$0(f'(0(f(0(f'(0(f(0(x)))))))))) \to 0(1(0(f'(0(f(0(x)))))))$$

$$0(f'(0(f(0(f'(0(f(0(x)))))))))) \to 0(f'(0(f(0(1(0(x)))))))$$

Along the lines of the proof of Prop. 58 we show *all* syntactical accidents are avoided. The lemma expresses an *invertibility* property (cf. [38, Thm. 8.4.20]): given the target term of a $\mathcal{T}^\omega$ reduction, the reduction can be reconstructed *up to permutation equivalence*.

▶ **Lemma 67.** *If γ, δ are co-initial and co-final $\mathcal{T}^\omega$ reductions, then $\gamma \simeq \delta$.*

▶ **Theorem 68.** *Developments and •-developments coincide in $[\mathcal{T}]^\omega$, if $\mathcal{T}$ is orthogonal.*

Proof. Since $\mathcal{T}$ is orthogonal by assumption, so is $[\mathcal{T}]$ by Lemma 61. Therefore, by Lemma 67: if $\gamma, \delta : a \twoheadrightarrow b$ are $[\mathcal{T}]^\omega$-reductions then $\gamma \simeq \delta$. It follows that if $\gamma : t \twoheadrightarrow\!\!\!\bullet\ t^\bullet$ and $\delta : t \twoheadrightarrow s$ for some $\epsilon : s \twoheadrightarrow t^\bullet$, then $\gamma \simeq \delta \cdot \epsilon$. Therefore, decomposing δ as $\delta_1 \cdot \phi \cdot \delta_2$ for some step $\phi : t' \to s'$, we have $\gamma/\delta_1 : t' \twoheadrightarrow\!\!\!\bullet\ s$ and $\phi \lesssim \gamma/\delta_1$, which by non-erasingness of $[\mathcal{T}]$ hence of $[\mathcal{T}]^\omega$ entails that ϕ is among the redex-patterns in γ/δ_1. Since this holds for each step, δ 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 $t^\bullet$. ◀

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** ───────

1 M. Abadi, L. Cardelli, P.-L. Curien, and J.-J. Lévy. Explicit substitutions. In F.E. Allen, editor, *Conference Record of the Seventeenth Annual ACM Symposium on Principles of Programming Languages, San Francisco, California, USA, January 1990*, pages 31–46. ACM Press, 1990. `doi:10.1145/96709.96712`.

2 P. Aczel. A general Church–Rosser theorem, 1978. corrections `http://www.ens-lyon.fr/LIP/REWRITING/MISC/AGRT_corrections.pdf`. URL: `http://www.ens-lyon.fr/LIP/REWRITING/MISC/AGeneralChurch-RosserTheorem.pdf`.

3 F. Baader and T. Nipkow. *Term Rewriting and All That.* Cambridge University Press, 1998.

4 H.P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics*, volume 103 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, 2nd revised edition, 1984.

5 A. Church and J.B. Rosser. Some properties of conversion. *Transactions of the American Mathematical Society*, 39:472–482, 1936. `doi:10.2307/1989762`.

6 P. Dehornoy. *Braids and Self-Distributivity*, volume 192 of *Progress in Mathematics*. Birkhäuser, 2000.

7 P. Dehornoy and V. van Oostrom. Z; proving confluence by monotonic single-step upper-bound functions. In *Logical Models of Reasoning and Computation (LMRC-08), Moscow*, 2008. URL: `http://cl-informatik.uibk.ac.at/users/vincent/research/publication/talk/lmrc060508.pdf`.

8 J. Endrullis, C. Grabmayer, D. Hendriks, J.W. Klop, and V. van Oostrom. Infinitary term rewriting for weakly orthogonal systems: Properties and counterexamples. *Logical Methods in Computer Science*, 10(2), 2014. `doi:10.2168/LMCS-10(2:7)2014`.

9 B. Felgenhauer, J. Nagele, V. van Oostrom, and C. Sternagel. The Z property. *Arch. Formal Proofs*, 2016, 2016. URL: `https://www.isa-afp.org/entries/Rewriting_Z.shtml`.

10 B. Felgenhaurer. Personal communication, 2017.

11 Y. Honda, K. Nakazawa, and K. Fujita. Confluence proofs of lambda-mu-calculi by Z theorem. *Studia Logica*, 2021. `doi:10.1007/s11225-020-09931-0`.

12 G. Huet. Confluent reductions: Abstract properties and applications to term rewriting systems. *J. ACM*, 27(4):797–821, October 1980. `doi:10.1145/322217.322230`.

13 G. Huet and J.-J. Lévy. Computations in orthogonal rewriting systems, Part I + II. In J.L. Lassez and G.D. Plotkin, editors, *Computational Logic – Essays in Honor of Alan Robinson*, pages 395–443, Cambridge MA, 1991. MIT Press. Update of: Call-by-need computations in non-ambiguous linear term rewriting systems, 1979.

14 J. Ketema, J.W. Klop, and V. van Oostrom. Vicious circles in rewriting systems. Technical Report E0427, Centrum voor Wiskunde en Informatica, December 2004. URL: `https://ir.cwi.nl/pub/11022/11022D.pdf`.

15 J.W. Klop, V. van Oostrom, and F. van Raamsdonk. Reduction strategies and acyclicity. In H. Comon-Lundh, C. Kirchner, and H. Kirchner, editors, *Rewriting, Computation and Proof, Essays Dedicated to Jean-Pierre Jouannaud on the Occasion of His* 60*th Birthday*, volume 4600 of *Lecture Notes in Computer Science*, pages 89–112. Springer, 2007. `doi:10.1007/978-3-540-73147-4_5`.

16 Y. Komori, N. Matsuda, and F. Yamakawa. A simplified proof of the Church–Rosser theorem. *Studia Logica: An International Journal for Symbolic Logic*, 102(1):175–183, 2014. `doi:10.1007/s11225-013-9470-y`.

17 J.-J. Lévy. *Réductions correctes et optimales dans le λ-calcul.* Thèse de doctorat d'état, Université Paris VII, 1978. URL: `http://pauillac.inria.fr/~levy/pubs/78phd.pdf`.

18 R. Loader. Notes on simply typed lambda calculus. ECS-LFCS- 98-381, Laboratory for Foundations of Computer Science, The University of Edinburgh, February 1998. URL: `http://www.lfcs.inf.ed.ac.uk/reports/98/ECS-LFCS-98-381/`.

19 L. Maranget. *La stratégie paresseuse.* Thèse de doctorat, Université Paris 7, 1992.

20   R. Mayr and T. Nipkow. Higher-order rewrite systems and their confluence. *Theor. Comput. Sci.*, 19(1):3–29, 1998. `doi:10.1016/S0304-3975(97)00143-6`.

21   P.-A. Melliès. Axiomatic rewriting theory VI residual theory revisited. In S. Tison, editor, *Rewriting Techniques and Applications, 13th International Conference, RTA 2002, Copenhagen, Denmark, July 22–24, 2002, Proceedings*, volume 2378 of *Lecture Notes in Computer Science*, pages 24–50. Springer, 2002. `doi:10.1007/3-540-45610-4_4`.

22   J. Nagele, V. van Oostrom, and C Sternagel. A short mechanized proof of the Church–Rosser theorem by the Z-property for the $\lambda\beta$-calculus in nominal Isabelle. In *5th International Workshop on Confluence, IWC 2016, Obergurgl, Austria, September 8–9, 2016, Online Proceedings*, 1016. URL: `http://www.csl.sri.com/users/tiwari/iwc2016/iwc2016.pdf`.

23   K. Nakazawa and K. Fujita. Compositional Z: Confluence proofs for permutative conversion. *Studia Logica: An International Journal for Symbolic Logic*, 104(6):1205–1224, 2016. `doi:10.1007/s11225-016-9673-0`.

24   R.P. Nederpelt. *Strong Normalization in a Typed Lambda Calculus with Lambda Structured Types*. PhD thesis, Technische Hogeschool Eindhoven, June 1973.

25   M.H.A. Newman. On theories with a combinatorial definition of "equivalence". *Annals of Mathematics*, 43:223–243, 1942. `doi:10.2307/2269299`.

26   T. Nipkow. Orthogonal higher-order rewrite systems are confluent. In M. Bezem and J.F. Groote, editors, *Typed Lambda Calculi and Applications, International Conference on Typed Lambda Calculi and Applications, TLCA '93, Utrecht, The Netherlands, March 16-18, 1993, Proceedings*, volume 664 of *Lecture Notes in Computer Science*, pages 306–317. Springer, 1993. `doi:10.1007/BFb0037114`.

27   V. van Oostrom. *Confluence for Abstract and Higher-Order Rewriting*. PhD thesis, Vrije Universiteit, Amsterdam, March 1994. URL: `https://research.vu.nl/files/62846778/complete%20dissertation.pdf`.

28   V. van Oostrom. Finite family developments. In H. Comon, editor, *Rewriting Techniques and Applications, 8th International Conference, RTA-97, Sitges, Spain, June 2-5, 1997, Proceedings*, volume 1232 of *Lecture Notes in Computer Science*, pages 308–322. Springer, 1997. `doi:10.1007/3-540-62950-5_80`.

29   V. van Oostrom. Reduce to the max, 1999. Unpublished manuscript. URL: `http://cl-informatik.uibk.ac.at/users/vincent/research/publication/pdf/max.pdf`.

30   V. van Oostrom. Random descent. In *RTA*, volume 4533 of *Lecture Notes in Computer Science*, pages 314–328. Springer, 2007. `doi:10.1007/978-3-540-73449-9_24`.

31   V. van Oostrom. Abstract rewriting. In A. Middeldorp, editor, *3rd International School on Rewriting, ISR 2008, Obergurgl, Austria, July 21–26, 2008*, 2008. Z-property in part 2 of the slides. URL: `http://cl-informatik.uibk.ac.at/isr-2008/html/b.4.html`.

32   V. van Oostrom. Confluence by decreasing diagrams; converted. In A. Voronkov, editor, *Rewriting Techniques and Applications, 19th International Conference, RTA 2008, Hagenberg, Austria, July 15-17, 2008, Proceedings*, volume 5117 of *Lecture Notes in Computer Science*, pages 306–320. Springer, 2008. `doi:10.1007/978-3-540-70590-1_21`.

33   V. van Oostrom and Y. Toyama. Normalisation by Random Descent. In *FSCD*, volume 52 of *LIPIcs*, pages 32:1–32:18. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2016. `doi:10.4230/LIPIcs.FSCD.2016.32`.

34   V. van Oostrom and Y. Venema. Term rewriting systems I and II. In *10th European Summer School in Logic, Language and Information, ESSLLI 98, Saarbrücken, Germany, August 17—28, 1998*, 2008. Course notes on braids in part I of the course. URL: `http://cl-informatik.uibk.ac.at/users/vincent/research/publication/pdf/braids.pdf`.

35   F. van Raamsdonk. *Confluence and Normalisation for Higher-Order Rewriting*. PhD thesis, Vrije Universiteit Amsterdam, May 1996. URL: `https://research.vu.nl/files/62847150/complete%20dissertation.pdf`.

36 R. Statman. There is no hyperrecurrent *s,k* combinator. Research Report 91-1332, Department of Mathematics, Carnegie Mellon University, Pittsburg, PA 15213, June 1991. URL: `http://shelf2.library.cmu.edu/Tech/53922203.pdf`.

37 M. Takahashi. Parallel reductions in λ-calculus. *Information and Computation*, 118:120–127, April 1995. `doi:10.1006/inco.1995.1057`.

38 Terese. *Term Rewriting Systems*. Cambridge University Press, 2003.

A Proofs omitted from the main text

Proofs of second and third items of Lem. 8. ▬ Assume $\to$ 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 $a \to b$, then $b \twoheadrightarrow a^\bullet \twoheadrightarrow b^\bullet$ by the Z-property for $\bullet$. Hence $a \twoheadrightarrow a^\bullet = a^\star$ by composition and definition of $\star$. If it is not, then $a \twoheadrightarrow a^\star = a$ by reflexivity and definition of $\star$.

To see that $\to$ has the Z-property for $\star$, suppose $a \to b$. By the Z-property for $\bullet$ and by definition of $\star$, then $b \twoheadrightarrow a^\bullet = a^\star \twoheadrightarrow b^\bullet$. 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 $b = a^\bullet = a^\star = b^\bullet$ 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 $\to$ has the Z-property for an extensive $\bullet$. To show $a \rightarrow\!\!\!\circ\!\!\!\rightarrow a^\bullet$, distinguish cases on whether there is some $\to$-step from a or not. If there is, say $a \to b$ then by the Z-property, $a \twoheadrightarrow b^\bullet \twoheadrightarrow a^\bullet$. If there is no $\to$-step from a, then extensivity of $\bullet$ entails $a = a^\bullet$. In either case, $a \twoheadrightarrow a^\bullet \twoheadrightarrow a^\bullet$ by reflexivity of $\twoheadrightarrow$, so $a \rightarrow\!\!\!\circ\!\!\!\rightarrow a^\bullet$ by definition of $\rightarrow\!\!\!\circ\!\!\!\rightarrow$. ◄

Proof of Lem. 34. (Sequentialisation) The proof is by induction on t. If t is a variable, then $ts = t\langle s]$ and we conclude by reflexivity. Otherwise, t has shape $t_1 t_2$ and we conclude using the IH to $ts = t_1 t_2 s \to t_1(t_2 s) \twoheadrightarrow t_1 t_2\langle s] = t\langle s]$ 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 $\twoheadrightarrow$. If t is a variable x, then $s \to s'$ and $t\langle s] = xs \to xs' = t\langle s']$ by compatibility of reduction. If $t = t_1 t_2$, we distinguish cases on where the step takes place:

▬ If the step takes place at the root of t, then $t = t_{11} t_{12} t_2 \to t_{11}(t_{12} t_2) = t'$ and we conclude by unfolding the definition of right-substitution twice on both sides to $t\langle s] = t_{11} t_{12} t_2 \langle s] \to t_{11}(t_{12} t_2 \langle s]) = t'\langle s]$;

▬ If the step takes place in t_1, then $t\langle s] = t_1 t_2 \langle s] \to t_1' t_2 \langle s] = t'\langle s]$ by compatibility of reduction;

▬ If the step takes place in t_2, then $t\langle s] = t_1 t_2 \langle s] \to t_1 t_2' \langle s] = t'\langle s]$ by the IH and compatibility of reduction;

▬ If the step takes place in s, then $t\langle s] = t_1 t_2 \langle s] \to t_1 t_2 \langle s'] = t\langle s']$ by the IH and compatibility of reduction.

(Substitution) The statement is shown by induction on t. If t is a variable, say x then $t\langle s]\langle r] = xs\langle r] = t\langle s\langle r]]$ by unfolding the definition of right-substitution. If t has shape $t_1 t_2$, then $t\langle s]\langle r] = t_1 t_2 \langle s]\langle r] = t_1 t_2 \langle s\langle r]] = t\langle s\langle r]]$ 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 $t = t^\bullet$ and we conclude by reflexivity of $\twoheadrightarrow$. Otherwise t has shape $t_1 t_2$, and we conclude by (Sequentialisation), the IH twice, (Compatible), and definition to $t_1 t_2 \twoheadrightarrow t_1 \langle t_2 \rangle \twoheadrightarrow t_1^\bullet \langle t_2^\bullet \rangle = (t_1 t_2)^\bullet$;

(Rhs)   By (Sequentialization) twice and (Substitution) we conclude $t^\bullet(s^\bullet r^\bullet) \twoheadrightarrow t^\bullet \langle s^\bullet \langle r^\bullet \rangle] = t^\bullet \langle s^\bullet] \langle r^\bullet] = (tsr)^\bullet$;

(Z)   As $\bullet$ maps to normal forms, we show a strengthening of the Z-property, $s \twoheadrightarrow t^\bullet = s^\bullet$, for all steps $t \to s$, 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 $t = (t_{11}t_{12})t_2 \to t_{11}(t_{12}t_2) = s$, and we conclude using (extensive), (Rhs), the definition, and (Substitution) to $t_{11}(t_{12}t_2) \twoheadrightarrow t_{11}^\bullet(t_{12}^\bullet t_2^\bullet) \twoheadrightarrow (t_{11}t_{12}t_2)^\bullet = t_{11}^\bullet \langle t_{12}^\bullet]\langle t_2^\bullet] = t_{11}^\bullet \langle t_{12}^\bullet \langle t_2^\bullet]]$;
- If the step takes place in t_1, say $t_1 \to s_1$, then we conclude using the IH, (Extensive), (Sequentialisation) and definition to $s_1 t_2 \twoheadrightarrow t_1^\bullet t_2^\bullet \twoheadrightarrow (t_1 t_2)^\bullet = t_1^\bullet \langle t_2^\bullet] = s_1^\bullet \langle t_2^\bullet] = (s_1 t_2)^\bullet$.
- If the step takes place in t_2 we proceed as in the previous item.   ◀

Proof of Lem. 40. Both items can we 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 $t_1 t_2 s \to t_1 s(t_2 s) \twoheadrightarrow t_1[s]t_2[s] = (t_1 t_2)[s]$, as $t_i s \twoheadrightarrow t_i[s]$ by the IH;

(Compatible) $t[s] = t^\sigma$ for the substitution σ mapping x to xs, and $t'[s'] = t'^{\sigma'}$ for σ' mapping x to xs'. Hence if $t \twoheadrightarrow t'$ and $s \twoheadrightarrow s'$ then $\sigma \twoheadrightarrow \sigma'$, hence $t^\sigma \twoheadrightarrow t'^{\sigma'}$ by compatibility of rewriting with substitution; and

(Substitution) For variables $x[s][r] = (xs)[r] = xrs[r] \twoheadrightarrow x[r][s[r]]$ by Sequentialisation twice, and for applications $(t_1 t_2)[s][r] = t_1[s][r]t_2[s][r] \twoheadrightarrow t_1[r][s[r]]t_2[r][s[r]] = (t_1 t_2)[r][s[r]]$ by the induction hypothesis twice.   ◀

Proof of Lem. 41. The items are proven by induction on t.

(Extensive) For variables $x = x^\bullet$, and for applications $ts \twoheadrightarrow t[s] \twoheadrightarrow t^\bullet[s^\bullet] = (ts)^\bullet$ 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 $tsr \to tr(sr)$. Then $tr(sr) \twoheadrightarrow t[r]s[r] = (ts)[r] \twoheadrightarrow (ts)^\bullet[r^\bullet] = (tsr)^\bullet$ by (Sequentialisation) and (Extensive), twice. Monotonicity of $\bullet$ holds by $(tsr)^\bullet = t^\bullet[s^\bullet][r^\bullet] \twoheadrightarrow t^\bullet[r^\bullet][s^\bullet[r^\bullet]] = (tr(sr))^\bullet$ using (Substitution).
- If $t_1 t_2 \to s_2 s_2$ because $t_i \to s_i$ and $t_{3-i} = s_{3-i}$ for some $i \in \{1, 2\}$, then $s_j \twoheadrightarrow t_j^\bullet \twoheadrightarrow s_j^\bullet$ for $j \in \{1, 2\}$, either by the IH, or (Extensive) and reflexivity. Using that, (Sequentialisation), and (Substitution) $s_1 s_2 \twoheadrightarrow s_1[s_2] \twoheadrightarrow t_1^\bullet[t_2^\bullet] = (t_1 t_2)^\bullet \twoheadrightarrow s_1^\bullet[s_2^\bullet] = (s_1 s_2)^\bullet$.   ◀

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]

---

[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.

The second part holds per construction of saturating left-hand sides of rules with memory.
◀

Proof of Lem. 65. For the first item first note that its only–if-direction requires $n > 0$ as otherwise $\mathcal{T}^n$ has *no* rules. Then, all (priming, labelling) operations for obtaining the rules of $\mathcal{T}^\omega$ from those of $\mathcal{T}$ are linear (only unary function symbols are added/removed) and redex-patterns overlapping in $\mathcal{T}^\omega$ still do so after removing labels and collapsing f'–f-pairs to f. $\mathcal{T}^n$ being a sub-system of $\mathcal{T}^\omega$ 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][21] stating that a term rewrite system is terminating iff r^σ is terminating for every rhs r of a rule and terminating substitution σ. 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 $f(\vec{t})$ with all $\vec{t}$ terminating by minimality. Hence an infinite reduction from it would have shape $f(\vec{t}) \twoheadrightarrow f(\vec{s}) = \ell^\sigma \to r^\sigma \twoheadrightarrow \ldots$ for some rule $\ell \to r$, substitution σ, and terms $\vec{s}$ such that $t_i \twoheadrightarrow s_i$ for all i. This is impossible as r^σ is terminating by assumption since σ is terminating as it assigns subterms of the $\vec{s}$ to variables[22] and each s_i is terminating as reduct of t_i.

To establish the assumption of the RHS lemma for $\mathcal{T}^n$ we prove the more general claim[23] that $(t^m)^\sigma$ *is terminating for every* $m \le n$, *term* t *over (primed) symbols in* $\mathcal{T}$, *and terminating substitution* σ. This suffices as per construction of $\mathcal{T}^n$ rhss of rules have this shape since labels in lhss are $< n$.[24] 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 σ is terminating, since the head symbol m is not affected by any step per construction of $\mathcal{T}^\omega$; labels occur in lhss only between (possibly primed) $\mathcal{T}$-symbols.

- Otherwise t has shape $f(\vec{t})$ for some (possibly primed) $\mathcal{T}$-symbol. Since each $(t_i^m)^\sigma$ 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

$$m(f((\vec{t}^m)^\sigma)) \twoheadrightarrow m(f(\vec{s})) = m(\ell^\tau) \to m((r^k)^\tau) \twoheadrightarrow \ldots$$

  for some $\mathcal{T}^n$ rule of shape $\ell \to r^k$ with k the maximum of the labels in ℓ plus one, substitution τ and r a term over (possibly primed) $\mathcal{T}$-symbols, and terms $\vec{s}$ such that

---

[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' $x \to x$, and for higher-order TRSs it would fail if non-pattern-lhss 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; rhss of rules are not closed (as rhss!) under subterms in general.

[24] Although *terms* of $\mathcal{T}^n$ 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 $\mathcal{T}^n$.

$(t_i^m)^\sigma \twoheadrightarrow s_i$ 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 $(t_i^m)^\sigma \twoheadrightarrow s_i$ guarantees that m is the head symbol of each s_i, and per construction of $\mathcal{T}^\omega$ 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 $\mathcal{T}$) below it. That τ is terminating follows from that it assigns subterms of the s_i to variables, which are reducts of the $(t_i^m)^\sigma$. ◀

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 $\mathcal{T}^n$ reductions $\gamma, \delta : t \twoheadrightarrow s$ we have $\gamma \simeq \delta$ by induction on t ordered by the union of $\leftarrow$ and the sub-term relation, well-founded since $\mathcal{T}^n$ is terminating. This suffices since any pair of $\mathcal{T}^\omega$-reductions is a pair of $\mathcal{T}^n$-reductions (take n greater than all labels occurring in the redex-patterns contracted in γ, δ), and $\mathcal{T}^n$-projection equivalence entails $\mathcal{T}^\omega$-projection equivalence.

Suppose γ, δ were minimal such that $\gamma \not\simeq \delta$. By residual theory, the peak γ, δ 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 γ', δ' must be non-empty. We may assume that γ, δ 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. γ, contains a head step. Since the system is orthogonal, [15, Lemma 1] yields then that δ does not contain a head step. Hence $\gamma/\delta \not\simeq \delta/\gamma$ since γ/δ contains a head step as projection of a reduction γ containing a head step over a reduction δ containing none, and δ/γ contains no head step as projection of δ containing none over another reduction γ, using that $\mathcal{T}^n$-rules are non-collapsing. Their sources being $\rightarrow^+$-reachable from t, $\gamma/\delta, \delta/\gamma$ contradicts minimality of γ, δ. ◀