1. The Parametric λ-Calculus

A calculus is a language equipped with some reduction rules. All the calculi we consider in this book share the same language, which is the language of λ-calculus, while they differ each other in their reduction rules. In order to treat them in an uniform way we define a parametric calculus, the $\lambda\Delta$-calculus, which gives rise to different calculi by different instantiations of the parameter Δ. In Part I we study the syntactical properties of the $\lambda\Delta$-calculus, and in particular those of its two most important instances, the call-by-name and the call-by-value λ-calculi. The $\lambda\Delta$-calculus has been introduced first in [85] and further studied in [74]. We use the terminology of [9].

1.1 The Language of λ-Terms

Definition 1.1.1 (The language Λ).
Let Var *be a countable set of variables. The set Λ of λ-terms is a set of words on the alphabet* $\mathrm{Var} \cup \{\ (\ ,\ )\ ,\ .\ ,\ \lambda\ \}$ *inductively defined as follows:*

- $x \in \mathrm{Var}$ *implies* $x \in \Lambda$,
- $M \in \Lambda$ *and* $x \in \mathrm{Var}$ *implies* $(\lambda x.M) \in \Lambda$ (abstraction),
- $M \in \Lambda$ *and* $N \in \Lambda$ *implies* $(MN) \in \Lambda$ (application).

λ-terms will be ranged over by Latin capital letters. Sets of λ-terms will be denoted by Greek capital letters.

Sometimes, we will refer to λ-terms simply as terms. The symbol $\equiv$ will denote syntactical identity of terms. We will use the following abbreviations in order to avoid an excessive number of parenthesis: $\lambda x_1...x_n.M$ will stand for $(\lambda x_1.(...(\lambda x_n.M)...))$, and $MN_1N_2...N_n$ will stand for $(...((MN_1)N_2)...N_n)$. Moreover, $\vec{M}$ will denote a sequence of terms $M_1,\ldots, M_n$ for some $n \geq 0$, and $\lambda\vec{x}.M$ and $\vec{M}\vec{N}$ will denote respectively $\lambda x_1\ldots x_n.M$ and $M_1\ldots M_m N_1\ldots N_n$ for some $n,m \geq 0$. The length of the sequence $\vec{N}$ is denoted by $\|\vec{N}\|$. By abusing the notation, $N \in \vec{N}$ denotes that the term N occurs in the sequence $\vec{N}$.

Example 1.1.2. $\lambda x.xx$, $\lambda x.x(\lambda z.zy)$, $\lambda y.(\lambda x.x)(\lambda uv.u)$ are examples of λ-terms. Some λ-terms have standard names for historical reasons. The names

that will be extensively used in this book are:
$I \equiv \lambda x.x$, $K \equiv \lambda xy.x$, $O \equiv \lambda xy.y$, $D \equiv \lambda x.xx$, $E \equiv \lambda xy.xy$.

Definition 1.1.3 (Subterms).
A term N is a subterm *of M if and only if one of the following conditions arises:*

- $M \equiv N$,
- $M \equiv \lambda x.M'$ *and N is a subterm of M',*
- $M \equiv PQ$ *and N is a subterm either of P or of Q.*

A term N occurs in a term M if and only if N is a subterm of M.

Example 1.1.4. The set of subterms of the term $\lambda x.x(\lambda z.zy)$ is

$$\{\lambda x.x(\lambda z.zy), x(\lambda z.zy), \lambda z.zy, zy, x, z, y\}.$$

The symbol "λ" plays the role of binder for variables, as formalized in the next definition.

Definition 1.1.5 (Free variables).

(i) *The set of* free variables *of a term M, denoted by* $\mathrm{FV}(M)$*, is inductively defined as follows:*
- $M \equiv x$ *implies* $\mathrm{FV}(M) = \{x\}$,
- $M \equiv \lambda x.M'$ *implies* $\mathrm{FV}(M) = \mathrm{FV}(M') - \{x\}$,
- $M \equiv PQ$ *implies* $\mathrm{FV}(M) = \mathrm{FV}(P) \cup \mathrm{FV}(Q)$.

A variable is bound *in M when it is not free in M.*

(ii) *A term M is* closed *if and only if* $\mathrm{FV}(M) = \emptyset$*. A term is* open *if it is not closed. For every subset of terms $\Theta \subseteq \Lambda$, we will denote with Θ^0 the restriction of Θ to closed terms.*

Example 1.1.6. $\mathrm{FV}\big(\lambda z.(\lambda x.x(\lambda z.zy))(\lambda xyz.yz)\big) = \{y\}$, $\mathrm{FV}\big(\lambda z.x(\lambda x.xy)\big) = \{x, y\}$, and $\mathrm{FV}\big((\lambda yx.x)y\big) = \{y\}$.

The replacement of a free variable by a term is the basic syntactical operation on Λ on which the definition of reduction rules will be based. But the replacement must respect the status of the variables: e.g., x can be replaced by $M \equiv \lambda y.zy$ in $\lambda u.xu$, so obtaining the term $\lambda u.(\lambda y.zy)u$, while the same replacement cannot take place in the term $\lambda z.xz$, since in the obtained term $\lambda z.(\lambda y.zy)z$ the free occurrence of z in M would become bound. The notion is formalized in the next definition.

Definition 1.1.7. *The statement* "M is free for x in N" *is defined by induction on N as follows:*

- *M is free for x in x;*
- *M is free for x in y;*
- *If M is free for x both in P and Q then M is free for x in PQ;*

- *If M is free for x in N and $x \not\equiv y$ and $y \notin \mathrm{FV}(M)$ then M is free for x in $\lambda y.N$.*

Example 1.1.8. $\lambda xy.xz$ is free for x and y in $(\lambda u.x)(\lambda u.xz)$ but is not free for u in both $\lambda xz.u$ and $\lambda zu.u$.

Let M be free for x in N; so $N[M/x]$ denotes the *simultaneous replacement* of all free occurrences of x in N by M. Clearly,

$$\mathrm{FV}(N[M/x]) = \begin{cases} \mathrm{FV}(N) & \text{if } x \notin \mathrm{FV}(N), \\ (\mathrm{FV}(N) - \{x\}) \cup \mathrm{FV}(M) & \text{otherwise.} \end{cases}$$

For example, $(\lambda x.u(xy))[xy/u]$ is not defined because xy is not free for u in $\lambda x.u(xy)$, while $(\lambda x.u(xu))[u(\lambda z.z)/u] \equiv \lambda x.u(\lambda z.z)(xu(\lambda z.z))$.

Let $\|\vec{N}\| = \|\vec{x}\|$; both $\vec{M}[N_1/x_1, ..., N_n/x_n]$ and $\vec{M}[\vec{N}/\vec{x}]$ are abbreviations for the simultaneous replacement of x_i by N_i in every M_j ($0 \leq i \leq \|\vec{x}\| = n$, $0 \leq j \leq \|\vec{M}\|$).

In the standard mathematical notation, the name of a bound variable is meaningless; for example, $\sum_{1 \leq i \leq n} i$ and $\sum_{1 \leq j \leq n} j$ both denote the sum of the first n natural numbers. Also in the language Λ, it is natural to consider the terms modulo names of bound variables. The renaming is formalized in the next definition.

Definition 1.1.9 (α-Reduction).

(i) $\lambda x.M \rightarrow_\alpha \lambda y.M[y/x]$ *if y is free for x in M and $y \notin \mathrm{FV}(M)$.*
(ii) $=_\alpha$ *is the reflexive, symmetric, transitive and contextual closure of* $\rightarrow_\alpha$.

Example 1.1.10. $\lambda x.x =_\alpha \lambda y.y =_\alpha \lambda z.z$, $\lambda xy.x =_\alpha \lambda xz.x$ and $\lambda xy.x =_\alpha \lambda yx.y$. On the other hand, $\lambda x.y \neq_\alpha \lambda x.x$ and $\lambda x.yx \neq_\alpha \lambda y.yy$.

In the entire book, we will consider terms modulo $=_\alpha$.

Thus we can also safely extend the notation $N[M/x]$ to the case where M is not free for x in N. In this case $N[M/x]$ denotes the result of replacing x by M in a term $N' =_\alpha N$ such that M is free for x in N'. Clearly such an N' always exists and the notation is well posed. So $(\lambda x.u(xy))[xy/u]$ is α-equivalent to the term $\lambda z.xy(zy)$.

An alternative way of denoting a simultaneous replacement is by explicitly using the notion of substitution. A *substitution* is a function from variables to terms. If $\mathbf{s}$ is a substitution and $\mathrm{FV}(M) = \{x_1, ..., x_n\}$, $\mathbf{s}(M)$ denotes $M[\mathbf{s}(x_1)/x_1, ..., \mathbf{s}(x_n)/x_n]$.

An important syntactical tool that will be extensively used in the following chapters is the notion of *context*. Informally, a context is a term that can contain some occurrences of a *hole* (denoted by the constant [.]) that can be filled by a term.

Definition 1.1.11 (Context).
Let Var *be a countable set of variables, and* $[.]$ *be a constant* (the hole)*.*

(i) *The set* Λ_C *of* contexts *is a set of words on* $\mathrm{Var} \cup \{\ (\ ,\ )\ ,\ .\ ,\ \lambda\ ,\ [.]\ \}$ *inductively defined as follows:*
- $[.] \in \Lambda_C$*;*
- $x \in \mathrm{Var}$ *implies* $x \in \Lambda_C$*;*
- $C[.] \in \Lambda_C$ *and* $x \in \mathrm{Var}$ *implies* $(\lambda x.C[.]) \in \Lambda_C$*;*
- $C_1[.] \in \Lambda_C$ *and* $C_2[.] \in \Lambda_C$ *implies* $(C_1[.]C_2[.]) \in \Lambda_C$*.*

  Contexts will be denoted by $C[.], C'[.], C_1[.], \ldots$ *.*

(ii) *A context of the shape:* $(\lambda\vec{x}.[.])\vec{P}$ *is a* head context*.*

(iii) *Let* $C[.]$ *be a context and* M *be a term. Then* $C[M]$ *denotes the term obtained by replacing by* M *every occurrence of* $[.]$ *in* $C[.]$*.*

We will use the same abbreviation notation for contexts that we used for terms.

Note that filling a hole in a context is *not* a substitution; in fact, free variables in M can become bound in $C[M]$. For example, filling the hole of $\lambda x.[.]$ with the free variable x results in the term $\lambda x.x$.

1.2 The $\boldsymbol{\lambda\Delta}$-Calculus

We will present some λ-calculi, all based on the language Λ, defined in the previous section, each one characterized by different reduction rules.

The $\lambda\Delta$-calculus is the language Λ equipped with a set $\Delta \subseteq \Lambda$ of input values, satisfying some closure conditions. Informally, input values represent partially evaluated terms that can be passed as parameters. Call-by-name and call-by-value parameter passing can be seen as the two most radical choices; parameters are not evaluated in the former policy, while in the latter they are evaluated until an output result is reached.

Most of the known variants of λ-calculus can be obtained from this parametric calculus by instantiating Δ in a suitable way. The set Δ of input values and the reduction $\to_\Delta$ induced by it are defined in Definition 1.2.1.

Definition 1.2.1. *Let* $\Delta \subseteq \Lambda$*.*

(i) *The* Δ-reduction *(*$\to_\Delta$*) is the contextual closure of the following rule:*

$$(\lambda x.M)N \to M[N/x] \textit{ if and only if } N \in \Delta.$$

$(\lambda x.M)N$ *is called a* Δ-redex *(or simply redex), and* $M[N/x]$ *is called its* Δ-contractum *(or simply contractum).*

(ii) $\to^*_\Delta$ *and* $=_\Delta$ *are respectively the reflexive and transitive closure of* $\to_\Delta$ *and the symmetric, reflexive and transitive closure of* $\to_\Delta$.

(iii) *A set* $\Delta \subseteq \Lambda$ *is said* set of input values *when the following conditions are satisfied:*

- $\mathrm{Var} \subseteq \Delta$ (Var-closure)*;*
- $P, Q \in \Delta$ *implies* $P[Q/x] \in \Delta$*, for each* $x \in \mathrm{Var}$ (substitution closure)*;*
- $M \in \Delta$ *and* $M \to_\Delta N$ *imply* $N \in \Delta$ (reduction closure).

(iv) *A term* is in Δ-normal form (Δ-*nf*) *if it has not* Δ*-redexes and it* has a Δ-normal form*, or it is* Δ-normalizing *if it reduces to a* Δ*-normal form; the set of* Δ*-nf is denoted by* Δ*-NF.*

(v) *A term is* Δ-strongly normalizing *if it is* Δ*-normalizing, and moreover there is not an infinite* Δ*-reduction sequence starting from it.*

The closure conditions on the set of input values need some comment. Since, as already said, input values represent partially evaluated terms, it is natural to ask that this partial evaluation is preserved by reduction, which is the rule on which the evaluation process is based. The substitution closure comes naturally from the fact that variables always belong to the set of input values.

In this book the symbol Δ will denote a generic set of input values. We will omit the prefix Δ in cases where it is clear from the context.

Example 1.2.2. Let I, K, O, D be the terms defined in the Example 1.1.2, and let M, N be input values. Then $IM \to_\Delta M$, so I has the behaviour of the identity function, $KMN \to^*_\Delta M$, $OMN \to^*_\Delta N$, $DM \to_\Delta MM$. If $D \in \Delta$ then $DD \to^*_\Delta DD$.

Now some possible sets of input values will be defined.

Definition 1.2.3. (i) $\Gamma = \mathrm{Var} \cup \{\lambda x.M \mid M \in \Lambda\}$.

(ii) Λ_I *is the language obtained from the grammar generating* Λ*, given in Definition 1.1.1, by modifying the formation rule for abstraction in the following way:*

$$(\lambda x.M) \in \Lambda_I \textit{ if and only if } M \in \Lambda \textit{ and } x \in \mathrm{Var} \textit{ and } x \textit{ occurs in } M.$$

The next property shows that there exists some set of input values, although not all sets of terms are sets of input values.

Property 1.2.4. (i) Λ is a set of input values.

(ii) Γ is a set of input values.

(iii) Λ_I is a set of input values.

(iv) Λ-NF is not a set of input values.

(v) $\mathrm{Var} \cup \Lambda\text{-NF}^0$ is a set of input values.

(vi) $\Upsilon = \mathrm{Var} \cup \{\lambda x.P \mid x \in \mathrm{FV}(P)\}$ is not a set of input values.

Proof. The first case is obvious. In cases 2, 3, and 5, it is easy to check that the closure properties of Definition 1.2.1 are satisfied. Λ-NF is not closed under substitution. It is easy to see that Υ is closed under substitution, but it is not closed under reduction. In fact, $\lambda x.KIx \in \Upsilon$, while $\lambda x.KIx \to_{\Upsilon} \lambda x.I \notin \Upsilon$. □

The choice $\Delta = \Lambda$ gives rise to the classical call-by-name λ-calculus [25], while $\Delta = \Gamma$ gives rise to a *pure* version (i.e. without constants) of the call-by-value λ-calculus, first defined by Plotkin [78].

The fact that $\mathrm{Var} \cup \Lambda\text{-NF}^0$ is a correct set of input values was first noticed in [39].

It is easy to check that every term M has the following shape:

$$\lambda x_1 ... x_n.\zeta M_1 ... M_m \qquad (n, m \geq 0),$$

where $M_i \in \Lambda$ are the *arguments* of M ($1 \leq i \leq m$) and ζ is the *head* of M. Here ζ is either a variable (*head variable*) or an application of the shape $(\lambda z.P)Q$, which can be either a redex (*head redex*) or not (*head block*), depending on the fact that Q belongs or not to the set Δ.

The natural interpretation of an abstraction term $\lambda x.M$ is a function whose formal parameter is x. The interpretation of an application $(\lambda x.M)N$, when $N \in \Delta$, is the application of the function $\lambda x.M$ to the actual parameter N, and so the Δ-reduction rule models the replacement of the formal parameter x by the actual parameter N in the body M of the function. Thus the Δ-normal form of a term, if it exists, can be seen as the final result of a computation.

The following fundamental theorem implies that this interpretation is correct, i.e. if the computation process stops, then the result is unique.

Theorem 1.2.5 (Confluence). [26, 74]
Let $M \to^*_\Delta N_1$ and $M \to^*_\Delta N_2$. There is Q such that both $N_1 \to^*_\Delta Q$ and $N_2 \to^*_\Delta Q$.

Proof. The proof is in Sect. 1.2.1. □

Corollary 1.2.6. *The Δ-normal form of a term, if it exists, is unique.*

Proof. Assume by absurdum that a term M has two different normal forms M_1 and M_2. Then, by the confluence theorem, there is a term N such that both M_1 and M_2 Δ-reduce to N, against the hypothesis that both are normal forms. □

It is natural to ask if the closure conditions on input values, given in Definition 1.2.1, are necessary in order to assure the confluence of the calculus. It can be observed that they are not strictly necessary, but a weaker version of them is needed.

Let $P \in \Delta$ be such that, for every $Q \not\equiv P$ such that $P \to^*_\Delta Q$, $Q \notin \Delta$. Thus $(\lambda x.M)P$ reduces both to $M[P/x]$ and to $(\lambda x.M)Q$, which do not have a common reduct, since the last term will be never a redex. Thus the weaker version of *reduction closure* that is necessary is the following: $M \in \Delta$ and $M \to^*_\Delta N$ imply that there is $P \in \Delta$ such that $N \to^*_\Delta P$.

On the other hand, let $N, P \in \Delta$ but for all Q such that $N[P/x] \to^*_\Delta Q$, $Q \notin \Delta$. Thus $(\lambda x.(\lambda y.M)N)P$ reduces both to $(\lambda y.M[P/x])N[P/x]$ and to $(M[N/y])[P/x]$, which do not have a common reduct. Thus the weaker version of the *substitution closure* that is necessary is the following: $P, Q \in \Delta$ implies there is $R \in \Delta$ such that $P[Q/x] \to^*_\Delta R$.

Assume $M \to^*_\Delta N$, and assume that there is more than one Δ-reduction sequence from M to N. The standardization theorem says that, in case the set of input values enjoys a particular property, there is a "standard" reduction sequence from M to N, reducing the redexes in a given order.

Let us introduce formally the notion of standard reduction sequence.

Definition 1.2.7. (i) *A symbol λ in a term M is* active *if and only if it is the first symbol of a Δ-redex of M.*

(ii) *The* Δ-sequentialization $(M)^\circ$ *of a term M is a function from Λ to Λ defined as follows:*
- $(xM_1...M_m)^\circ = x(M_1)^\circ...(M_m)^\circ$*;*
- $((\lambda x.P)QM_1...M_m)^\circ = (\lambda x.P)^\circ(Q)^\circ(M_1)^\circ...(M_m)^\circ$*, if $Q \in \Delta$;*
- $((\lambda x.P)QM_1...M_m)^\circ = (Q)^\circ(\lambda x.P)^\circ(M_1)^\circ...(M_m)^\circ$*, if $Q \notin \Delta$;*
- $(\lambda x.P)^\circ = \lambda x.(P)^\circ$.

(iii) *The* degree *of a redex R in M is the numbers of λ's that both are active in M and occur on the left of $(R)^\circ$ in $(M)^\circ$.*

(iv) *The* principal redex *of M, if it exists, is the redex of M with minimum degree. The* principal reduction $M \to^p_\Delta N$ *denotes that N is obtained from M by reducing the principal redex of M. Moreover, $\to^{*p}_\Delta$ is the reflexive and transitive closure of $\to^p_\Delta$.*

(v) *A sequence $M \equiv P_0 \to_\Delta P_1 \to_\Delta ... \to_\Delta P_n \to_\Delta N$ is* standard *if and only if the degree of the redex contracted in P_i is less than or equal to the degree of the redex contracted in P_{i+1}, for every $i < n$.*

We denote by $M \to^\circ_\Delta N$ a standard reduction sequence from M to N.

It is important to notice that the degree of a redex can change during the reduction; in particular, the redex of minimum degree always has degree zero. Moreover, note that the reduction sequences of length 0 and 1 are always standard. It is easy to check that, for every M, the Λ-sequentialization is $(M)^\circ \equiv M$; thus in this case the redex of degree 0 is always the leftmost one.

Example 1.2.8. (i) Let $\Delta = \Lambda$, and let $M \equiv (\lambda x.x(KI))(II)$. Thus M has degree 0, KI has degree 1 and II has degree 2 (in the term M). The following reduction sequence is standard:
$(\lambda x.x(KI))(II) \to_\Lambda (II)(KI) \to_\Lambda I(KI) \to_\Lambda I(\lambda y.I)$.

(ii) Let M be as before, and let $\Delta = \Gamma$. Thus II has degree 0, and KI has degree 1. Note that now M is no more a redex. The following reduction sequence is standard:
$(\lambda x.x(KI))(II) \to_\Gamma (\lambda x.x(KI))I \to_\Gamma I(KI) \to_\Gamma I(\lambda y.I) \to_\Gamma \lambda y.I$.

(iii) Let M be as before, and let $\Delta = \mathrm{Var} \cup \Lambda\text{-NF}^0$. Thus KI has degree 0 and II has degree 1. Also in this case M is not a redex. The following reduction sequence is standard:
$(\lambda x.x(KI))(II) \to_\Delta (\lambda x.x(KI))I \to_\Delta (\lambda x.x(\lambda y.I))I$.

The notion of a standard set of input values, which is given in Definition 1.2.9 is key for having the standardization property.

Definition 1.2.9 (Standard input values).
A set Δ of input values is standard *if and only if $M \notin \Delta$ and $M \to^*_\Delta N$ by reducing at every step a not principal redex imply $N \notin \Delta$.*

Now the standardization property can be stated.

Theorem 1.2.10 (Standardization). [74]
Let Δ be standard. $M \to^*_\Delta N$ implies there is a standard reduction sequence from M to N.

Proof. The proof is in Sect. 1.2.1. □

The next property shows that some sets of input values are standard, while some are not standard.

Property 1.2.11. (i) Λ and Γ are standard.
(ii) $\mathrm{Var} \cup \Delta\text{-NF}^0$ is standard, for every Δ.
(iii) Λ_I is not standard.

Proof. (i) Λ is trivially standard. Let us consider Γ; we will prove that, if $M \notin \Gamma$, and $M \to_\Gamma N$ through a not principal reduction, then $N \notin \Gamma$.
$M \notin \Gamma$ implies that M has one of the following shapes:
1. $yM_1...M_m$ $(m > 1)$.
2. $(\lambda x.M_1)M_2...M_m$ $(m \geq 2)$ and either $(\lambda x.M_1)M_2$ is a redex or it is a head block.

Case 1 is trivial, since M can never be reduced to a term in Γ.
In case 2, if $M_2 \in \Gamma$ then the principal redex is $(\lambda x.M_1)M_2$, while if $M_2 \notin \Gamma$ then if $M_2 \notin \Gamma\text{-NF}$ the principal redex is in M_2; if $M_2 \in \Gamma\text{-NF}$ then the principal redex is in some M_j $(j \leq 3)$. So the reduction of a not principal redex cannot produce a term belonging to Γ.

(ii) $\mathrm{Var} \cup \Delta\text{-NF}^0$ is standard since not principal reductions preserve the presence of the redex of minimum degree.

(iii) Consider the term, $M \equiv \lambda x.x(DD)((\lambda z.I)I)$. Clearly $M \notin \Lambda_I$ and the principal redex of M is DD. So $M \to_{\Lambda_I} \lambda x.x(DD)I \in \Lambda_I$ and in this reduction the reduced redex is not principal, while for every sequence of $\to^{*p}_{\Lambda_I}$ reductions; $M \to^{*p}_{\Lambda_I} M \notin \Lambda_I$. □

It is easy to see that the substitution closure on input values, given in Definition 1.2.1, is necessary in order to assure the standardization property.

In fact, let $M, N \in \Delta$ and $M[N/x] \notin \Delta$. The following non-standard reduction sequence $(\lambda x.IM)N \rightarrow_\Delta (\lambda x.M)N \rightarrow_\Delta M[N/x]$ does not have a standard counterpart, in fact $I(M[N/x]) \not\rightarrow_\Delta M[N/x]$.

Theorem 1.2.12. *The condition that Δ is standard is necessary and sufficient for the $\lambda\Delta$-calculus enjoys the standardization property.*

Proof. The sufficiency of the condition is a consequence of the Standardization Theorem. To prove its necessity, assume Δ is not standard; we can find a term $M \notin \Delta$ such that $M \rightarrow^*_\Delta N \in \Delta$, without reducing the principal redex. Hence $IM \rightarrow_\Delta IN \rightarrow_\Delta N$, by reducing first a redex of degree different from 0 and then a redex of degree 0. Clearly, there is no way of commuting the order of reductions. □

An important consequence of the standardization property is the fact that the reduction sequence reducing, at every step, the principal redex is normalizing, as shown in Corollary 1.2.13.

Corollary 1.2.13. *Let Δ be standard.*
If $M \rightarrow^*_\Delta N$ and N is a normal form then $M \rightarrow^{*p}_\Delta N$.

Proof. By Corollary 1.2.6 and by the definition of the standard set of input values. □

Example 1.2.14. (i) Let $\Delta = \Lambda$. The term $KI(DD)$ has Λ-normal form I. In fact, the principal Λ-reduction sequence is $KI(DD) \rightarrow_\Lambda (\lambda y.I)(DD) \rightarrow_\Lambda I$, while the Λ-reduction sequence choosing at every step the rightmost Λ-redex never stops. Notice that, if we choose $\Delta = \Gamma$, $KI(DD)$ has not Γ-normal form.

(ii) The term $II(II(II))$ is Λ-strongly normalizing and Γ-strongly normalizing, while $KI(DD)$ is neither Λ-strongly normalizing nor Γ-strongly normalizing.

(iii) Let $\Delta = \text{Var} \cup \Lambda\text{-NF}^0$. Thus $I(K(xx))$ is the Δ-normal form of term $I(II)(K(xx))$.

Remark 1.2.15. The first notion of standardization was given, for the $\lambda\Lambda$-calculus, by Curry and Feys [34, 35]. With respect to their notion, if $M \rightarrow^*_\Lambda N$ then there is a standard reduction sequence from M to N, but this reduction sequence is not necessarily unique. For instance, $\lambda x.x(II)(II) \rightarrow_\Lambda \lambda x.xI(II) \rightarrow_\Lambda \lambda x.xII$ and $\lambda x.x(II)(II) \rightarrow_\Lambda \lambda x.x(II)I \rightarrow_\Lambda \lambda x.xII$ are both standard reduction sequences. Klop [58] introduced a notion of strong standardization, according to which, if $M \rightarrow^*_\Lambda N$, then there is a unique strongly standard reduction sequence from M to N, and he designed an algorithm for transforming a reduction sequence into a strongly standard one. According to his notion, in the example before only the first reduction sequence is standard. Our definition, when restricted to the $\lambda\Lambda$-calculus, is quite similar to the strong standardization. In fact, according to our definition, the standard reduction sequence is unique, but in some degenerated case: e.g. for $\Delta = \Lambda$, there are infinite reduction sequences from $x(DD)$ to $x(DD)$, each one performing a different number of Λ-reductions.

Plotkin [78] extended the notion of standardization to the $\lambda\Gamma$-calculus. His notion of standardization is not strong using Klop's terminology. Our definition, when restricted to $\lambda\Gamma$-calculus, is similar to a strong version of Plotkin's standardization. The advantage of our notion of standardization is the validity of Corollary 1.2.13, i.e. the fact that the principal reduction is Δ-normalizing.

A notion that will play an important role in what follows is that one of solvability.

Definition 1.2.16. (i) *An head context* $(\lambda\vec{x}[.])\vec{P}$ *is* Δ-valuable *if and only if each* $P \in \vec{P}$ *is such that* $P \in \Delta$.
(ii) *A term* M *is* Δ-solvable *if and only if there is a* Δ*-valuable* head context $C[.] \equiv (\lambda\vec{x}.[.])\vec{N}$ *such that:*

$$C[M] =_{\Delta} I.$$

(iii) *A term is* Δ-unsolvable *if and only if it is not* Δ*-solvable.*

Note that $(\lambda\vec{x}.[.])\vec{N} =_{\Delta} I$ means $(\lambda\vec{x}.[.])\vec{N} \to^*_{\Delta} I$, since I is in Δ-nf, for every Δ.

We will abbreviate Δ-solvable and Δ-unsolvable respectively as solvable and unsolvable, when the meaning is clear from the context. Informally speaking, a solvable term is a term that is in some sense *computationally meaningful*. In fact, let $M \in \Lambda^0$ be solvable, and let P be an input value; we can always find a sequence $\vec{N}$ of terms such that $M\vec{N}$ reduces to P: just take the sequence $\vec{Q}$ such that $M\vec{Q} =_{\Delta} I$, which exists since M is solvable, and pose $\vec{N} \equiv \vec{Q}P$. So a closed solvable term can mimic the behaviour of any term, if applied to suitable arguments.

It would be interesting to syntactically characterize the solvable terms. Unfortunately, there is not a general characterization for the $\lambda\Delta$-calculus, so we will study this problem for some particular instances of Δ.

Example 1.2.17. (i) Consider the two sets of input values Λ and Γ. In both calculi, the term I is solvable, while DD is unsolvable. $\lambda x.x(DD)$ is an example of a term that is Λ-solvable and Γ-unsolvable. In fact, $(\lambda x.x(DD))O \to^*_{\Lambda} I$, while there is no term P such that $P(DD) \to^*_{\Gamma} I$, since $DD \notin \Gamma$ and $DD \to^*_{\Gamma} DD$.
(ii) Let Φ be the set of input values $\mathrm{Var} \cup \Lambda\text{-NF}^0$. Then $I(\lambda x.I(xx)) \in \Phi$-NF is a Φ-unsolvable term.

In order to understand the behaviour of unsolvable terms, it is important to stress some of their closure properties.

Property 1.2.18. (i) The unsolvability is preserved by substitution of variables by input values.
(ii) The unsolvability is preserved by Δ-valuable head contexts.

Proof. Let M be unsolvable.

(i) By contraposition let us assume $M[P/z]$ to be solvable for some input values P. Then there is a Δ-valuable head context $C[.] \equiv (\lambda\vec{x}.[.])\vec{Q}$ such that $C[M[P/z]] \to^*_\Delta I$.
Without loss of generality, we can assume $\|\vec{Q}\| > \|\vec{x}\|$. Indeed, in the case $\|\vec{Q}\| \le \|\vec{x}\|$, we can choose a closed solvable term N such that there is $\vec{R}$ such that $N\vec{R} \to^*_\Delta I$ and $\|\vec{R}\| = \|\vec{x}\| - \|\vec{Q}\|$, and then consider the Δ-valuable context $C[.]N\vec{R}$. So let $\vec{Q} \equiv \vec{Q_1}\vec{Q_2}$, where $\|\vec{Q_1}\| = \|\vec{x}\|$.
$(\lambda\vec{x}.M[P/z])\vec{Q_1}\vec{Q_2} \to^*_\Delta I$ implies $(\lambda\vec{x}.(\lambda z.M)P)\vec{Q_1}\vec{Q_2} \to^*_\Delta I$ (since $P \in \Delta$). This in turn implies $(\lambda z.(\lambda\vec{x}.M)\vec{Q_1})(P[\vec{Q_1}/\vec{x}])\vec{Q_2} \to^*_\Delta I$ and $(\lambda z\vec{x}.M)(P[\vec{Q_1}/\vec{x}])\vec{Q_1}\vec{Q_2} \to^*_\Delta I$, because by α-equivalence we can assume $z \notin \mathrm{FV}(Q_1)$ and $z \notin \vec{x}$. But $P[\vec{Q_1}/\vec{x}] \in \Delta$ (since input values are closed under substitution) which means that the Δ-valuable head context $C'[.] \equiv (\lambda z\vec{x}.[.])(P[\vec{Q_1}/\vec{x}])\vec{Q_1}\vec{Q_2}$ is such that $C'[M] \to^*_\Delta I$.

(ii) By contraposition let us assume $C'[M]$ to be solvable for some Δ-valuable head context $C'[.] \equiv (\lambda\vec{z}.[.])\vec{P}$. Then there is a Δ-valuable head context $C[.] \equiv (\lambda\vec{x}.[.])\vec{Q}$, such that $C[C'[M]] \to^*_\Delta I$. If $\vec{z} \equiv \vec{z_0}\vec{z_1}$ and $\|\vec{P}\| = \|\vec{z_0}\|$ then $C[C'[M]] \to^*_\Delta C[\lambda\vec{z_1}.M[\vec{P}/\vec{z_0}]] \to^*_\Delta I$, thus $M[\vec{P}/\vec{z_0}]$ is solvable, and by the previous part of this property M is also solvable. Otherwise $\vec{P} \equiv \vec{P_0}\vec{P_1}$, $\|\vec{P_1}\| > 1$ and $\|\vec{P_0}\| = \|\vec{z}\|$. Thus

$$C[C'[M]] \to^*_\Delta C[M[\vec{P_0}/\vec{z}]\vec{P_1}] \equiv (\lambda\vec{x}.M[\vec{P_0}/\vec{z}]\vec{P_1})\vec{Q} \to^*_\Delta I.$$

Without loss of generality we can assume $\|\vec{Q}\| > \|\vec{x}\|$, $\vec{Q} \equiv \vec{Q_0}\vec{Q_1}$ and $\|\vec{Q_0}\| = \|\vec{x}\|$. So

$$\begin{aligned}(\lambda\vec{x}.M[\vec{P_0}/\vec{z}]\vec{P_1})\vec{Q} \to^*_\Delta (M[\vec{P_0}/\vec{z}]\vec{P_1})[\vec{Q_0}/\vec{x}]\vec{Q_1} \equiv \\ (M[\vec{P_0}/\vec{z}][\vec{Q_0}/\vec{x}])(\vec{P_1}[\vec{Q_0}/\vec{x}])\vec{Q_1} \to^*_\Delta I,\end{aligned}$$

which implies $(M[\vec{P_0}/\vec{z}][\vec{Q_0}/\vec{x}])$ solvable. Again the proof follows from part (i) of this property. □

We will see that in all the calculi we will study in the following, the property of solvability is not preserved by either substitution or by head contexts. As an example in the $\lambda\Lambda$-calculus xD is Λ-solvable, but $xD[D/x]$ is not Λ-solvable.

1.2.1 Proof of Confluence and Standardization Theorems

Both the proofs are based on the notion of parallel reduction.

Definition 1.2.19. *Let Δ be a set of input values.*

(i) *The* deterministic parallel reduction $\hookrightarrow_\Delta$ *is inductively defined as follows:*
   1. $x \hookrightarrow_\Delta x$*;*
   2. $M \hookrightarrow_\Delta N$ *implies* $\lambda x.M \hookrightarrow_\Delta \lambda x.N$*;*
   3. $M \hookrightarrow_\Delta M'$*,*$N \hookrightarrow_\Delta N'$ *and* $N \in \Delta$ *imply* $(\lambda x.M)N \hookrightarrow_\Delta M'[N'/x]$*;*
   4. $M \hookrightarrow_\Delta M'$*,*$N \hookrightarrow_\Delta N'$ *and* $N \notin \Delta$ *imply* $MN \hookrightarrow_\Delta M'N'$*.*

(ii) *The* nondeterministic parallel reduction $\Rightarrow_\Delta$ *is inductively defined as follows:*
   1. $x \Rightarrow_\Delta x$*;*
   2. $M \Rightarrow_\Delta N$ *implies* $\lambda x.M \Rightarrow_\Delta \lambda x.N$*;*
   3. $M \Rightarrow_\Delta M'$*,*$N \Rightarrow_\Delta N'$ *and* $N \in \Delta$ *imply* $(\lambda x.M)N \Rightarrow_\Delta M'[N'/x]$*;*
   4. $M \Rightarrow_\Delta M'$*,*$N \Rightarrow_\Delta N'$ *imply* $MN \Rightarrow_\Delta M'N'$*.*

Roughly speaking, the deterministic parallel reduction reduces in one step all the redexes present in a term, while the nondeterministic one reduces a subset of them.

Example 1.2.20. Let $M \equiv I(II)$. If $\Delta \equiv \Lambda$ then $M \hookrightarrow_\Delta I$, while $M \Rightarrow_\Delta M$, $M \Rightarrow_\Delta II$ and $M \Rightarrow_\Delta I$. If $\Delta \equiv \Gamma$ then $M \hookrightarrow_\Delta II$ while $M \Rightarrow_\Delta M$ and $M \Rightarrow_\Delta II$.

The following lemma shows the relation between the $\Rightarrow_\Delta$ and $\to_\Delta$ reductions.

Lemma 1.2.21. *Let Δ be a set of input values.*

(i) $M \to_\Delta N$ *implies* $M \Rightarrow_\Delta N$*.*
(ii) $M \Rightarrow_\Delta N$ *implies* $M \to_\Delta^* N$*.*
(iii) $\to_\Delta^*$ *is the transitive closure of* $\Rightarrow_\Delta$*.*

Proof. Easy. □

$\Rightarrow_\Delta$ enjoys a useful substitution property.

Lemma 1.2.22. *Let $M \Rightarrow_\Delta M'$ and $N \Rightarrow_\Delta N'$.*
If $N \in \Delta$ then $M[N/x] \Rightarrow_\Delta M'[N'/x]$.

Proof. By induction on M. Let us prove just the most difficult case, i.e. the term M is a Δ-redex. Let $M \equiv (\lambda z.P)Q$, $Q \in \Delta$, $P \Rightarrow_\Delta P'$, $Q \Rightarrow_\Delta Q'$ and $M' \equiv P'[Q'/z]$. By induction $P[N/x] \Rightarrow_\Delta P'[N'/x]$ and $Q[N/x] \Rightarrow_\Delta Q'[N'/x]$, where $Q'[N'/x] \in \Delta$ for the closure conditions on Δ. Thus

$$((\lambda z.P)Q)[N/x] \equiv (\lambda z.P[N/x])Q[N/x] \Rightarrow_\Delta$$

$$P'[N'/x][Q'[N'/x]/z] \equiv (P'[Q'/z])[N'/x]$$

by point 3 of the definition of $\Rightarrow_\Delta$. □

The next property, whose proof is obvious, states that, for every term M, there is a unique term N such that $M \hookrightarrow_{\Delta} N$.

Property 1.2.23. $M \hookrightarrow_{\Delta} P$ and $M \hookrightarrow_{\Delta} Q$ implies $P \equiv Q$.

Proof. Trivial. □

Let $[M]_{\Delta}$ be the term such $M \hookrightarrow_{\Delta} [M]_{\Delta}$. In the literature $[M]_{\Delta}$ is called the *complete development* of M (see [93]). The following lemma holds.

Lemma 1.2.24. *$M \Rightarrow_{\Delta} N$ implies $N \Rightarrow_{\Delta} [M]_{\Delta}$.*

Proof. By induction on M.

- If $M \equiv x$, then $N \equiv x$ and $[M]_{\Delta} \equiv x$.
- If $M \equiv \lambda x.P$ then $N \equiv \lambda x.Q$, for some Q such that $P \Rightarrow_{\Delta} Q$. By induction $Q \Rightarrow_{\Delta} [P]_{\Delta}$, and so $N \Rightarrow_{\Delta} \lambda x.[P]_{\Delta} \equiv [M]_{\Delta}$.
- If $M \equiv P_1P_2$ and it is not a Δ-redex, then $N \equiv Q_1Q_2$ for some Q_1 and Q_2 such that $P_1 \Rightarrow_{\Delta} Q_1$ and $P_2 \Rightarrow_{\Delta} Q_2$. So, by induction, $Q_1 \Rightarrow_{\Delta} [P_1]_{\Delta}$ and $Q_2 \Rightarrow_{\Delta} [P_2]_{\Delta}$, which implies $N \Rightarrow_{\Delta} [P_1]_{\Delta}[P_2]_{\Delta} \equiv [M]_{\Delta}$.
- If $M \equiv (\lambda x.P_1)P_2$ is a redex (i.e. $P_2 \in \Delta$) then either $N \equiv (\lambda x.Q_1)Q_2$ or $N \equiv Q_1[Q_2/x]$, for some Q_i such that $P_i \Rightarrow_{\Delta} Q_i$ $(1 \leq i \leq 2)$. By induction, $Q_i \Rightarrow_{\Delta} [P_i]_{\Delta}$ $(1 \leq i \leq 2)$. Note that $[P_2]_{\Delta} \in \Delta$ by Lemma 1.2.21.(ii). In both cases, $N \Rightarrow_{\Delta} [P_1]_{\Delta}[[P_2]_{\Delta}/x] \equiv [M]_{\Delta}$, in the former case simply by induction, and in the latter both by induction and by Lemma 1.2.22. □

$$\begin{tikzcd}
 & M \arrow[dl, Rightarrow, "\Delta"'] \arrow[dr, Rightarrow, "\Delta"] & \\
N_0 \arrow[dr, Rightarrow, "\Delta"'] & & N_1 \arrow[dl, Rightarrow, "\Delta"] \\
 & N_2 &
\end{tikzcd}$$

Fig. 1.1. Diamond property.

The proof of confluence follows the Takahashi pattern [93], which is a simplification of the original proof made by Taït and Martin Löf for classical $\lambda\Lambda$-calculus. It is based on the property that a reduction that is the transitive closure of another one enjoying the Diamond Property is confluent.

Lemma 1.2.25 (Diamond property of $\Rightarrow_{\Delta}$).
If $M \Rightarrow_{\Delta} N_0$ and $M \Rightarrow_{\Delta} N_1$ then there is N_2 such that both $N_0 \Rightarrow_{\Delta} N_2$ and $N_1 \Rightarrow_{\Delta} N_2$.

Proof. By Lemma 1.2.24, $M \Rightarrow_\Delta N$ implies $N \Rightarrow_\Delta [M]_\Delta$. So, if $M \Rightarrow_\Delta M_1$ and $M \Rightarrow_\Delta M_2$, then both $M_1 \Rightarrow_\Delta [M]_\Delta$ and $M_2 \Rightarrow_\Delta [M]_\Delta$, as shown in Fig. 1.1 (pag. 15). □

$$
\begin{tikzcd}
M \arrow[r, Rightarrow, "\Delta"'] \arrow[d, Rightarrow, "\Delta"] \arrow[rrrr, bend left, "\Delta^*"] \arrow[dddd, bend right, "\Delta^*"'] & N_0^1 \arrow[r, Rightarrow, "\Delta"'] \arrow[d, Rightarrow, "\Delta"] & \cdots \arrow[r, Rightarrow, "\Delta"'] & N_0^{n_0} \arrow[r, Rightarrow, "\Delta"'] \arrow[d, Rightarrow, "\Delta"] & N_0 \arrow[d, Rightarrow, "\Delta"] \\
N_1^1 \arrow[r, Rightarrow, "\Delta"'] \arrow[d, Rightarrow, "\Delta"] & {[M^1]_\Delta} \arrow[r, Rightarrow, "\Delta"'] \arrow[d, Rightarrow, "\Delta"] & \cdots \arrow[r, Rightarrow, "\Delta"'] & \cdots \arrow[r, Rightarrow, "\Delta"'] & \vdots \arrow[d, Rightarrow, "\Delta"] \\
\vdots \arrow[d, Rightarrow, "\Delta"] & \vdots \arrow[d, Rightarrow, "\Delta"] & {} \arrow[dr, dotted, "\Delta^*"] & & \vdots \arrow[d, Rightarrow, "\Delta"] \\
N_1^{n_1} \arrow[r, Rightarrow, "\Delta"'] \arrow[d, Rightarrow, "\Delta"] & \vdots \arrow[d, Rightarrow, "\Delta"] & & {} & \vdots \arrow[d, Rightarrow, "\Delta"] \\
N_1 \arrow[r, Rightarrow, "\Delta"'] & \cdots \arrow[r, Rightarrow, "\Delta"'] & \cdots \arrow[r, Rightarrow, "\Delta"'] & \cdots \arrow[r, Rightarrow, "\Delta"'] & N_2
\end{tikzcd}
$$

Fig. 1.2. Diamond closure.

▼ **Proof of Confluence Theorem** (Theorem 1.2.5 pag. 8).

By Property 1.2.21.(iii), $\rightarrow^*_\Delta$ is the transitive closure of $\Rightarrow_\Delta$. This means that there are $N_0^1, ..., N_0^{n_0}$, $N_1^1, ..., N_1^{n_1}$ $(n_0, n_1 \geq 1)$ such that $M \Rightarrow_\Delta N_0^1 ... \Rightarrow_\Delta N_0^{n_0} \Rightarrow_\Delta N_0$ and $M \Rightarrow_\Delta N_1^1 ... \Rightarrow_\Delta N_m^{n_1} \Rightarrow_\Delta N_1$. Then the proof follows by repeatedly applying the diamond property of $\Rightarrow_\Delta$ (diamond closure), as shown in Fig. 1.2. ■

The rest of this subsection is devoted to the proof of the standardization theorem. First, we need to establish some technical results.

Let $M \Rightarrow^\circ_\Delta N$ denote "$M \rightarrow^\circ_\Delta N$ and $M \Rightarrow_\Delta N$".
The following lemma, at the point (ii), shows that a nondeterministic parallel reduction can always be transformed into a standard reduction sequence.

Lemma 1.2.26. *Let $\vec{P}, \vec{Q}$ be two sequences of terms such that $\|\vec{P}\| = \|\vec{Q}\|$; moreover, let $P_i \in \Delta$ and $P_i \Rightarrow^\circ_\Delta Q_i$ for all $i \leq \|\vec{P}\|$.*

(i) *If $M \Rightarrow^\circ_\Delta N$ then $M[\vec{P}/\vec{x}] \Rightarrow^\circ_\Delta N[\vec{Q}/\vec{x}]$.*
(ii) *If $M \Rightarrow_\Delta N$ then $M \Rightarrow^\circ_\Delta N$.*

Proof. Parts (i) and (ii) can be proved by mutual induction on M.

(i) By Lemma 1.2.22, $M[\vec{P}/\vec{x}] \Rightarrow_\Delta N[\vec{Q}/\vec{x}]$, hence it suffices to show that $M[\vec{P}/\vec{x}] \rightarrow^\circ_\Delta N[\vec{Q}/\vec{x}]$.

Let $M \equiv \lambda y_1...y_h.\zeta M_1...M_m$ $(h, m \in \mathbb{N})$, where either ζ is a variable or $\zeta \equiv (\lambda z.T)U$.

If $h > 0$, then the proof follows by induction.

Let $h = 0$, thus $N \equiv \xi N_1...N_m$ such that $\zeta \Rightarrow^\circ_\Delta \xi$ and $M_i \Rightarrow^\circ_\Delta N_i$; furthermore, let $M'_i \equiv M_i[\vec{P}/\vec{x}]$ and $N'_i \equiv N_i[\vec{Q}/\vec{x}]$ $(1 \leq i \leq m)$.

The proof is organized according to the possible shapes of ζ.

1. Let ζ be a variable. If $m = 0$ then the proof is trivial, so let $m > 0$. There are two cases to be considered.

1.1. $\zeta \notin \vec{x}$, so $\xi[\vec{Q}/\vec{x}] \equiv \zeta$. By induction $M_i[\vec{P}/\vec{x}] \rightarrow^\circ_\Delta N_i[\vec{Q}/\vec{x}]$ and the standard reduction sequence is

$$\zeta M'_1...M'_m \rightarrow^\circ_\Delta \zeta N'_1 M'_2...M'_m \rightarrow^\circ_\Delta \; ..... \; \rightarrow^\circ_\Delta \zeta N'_1...N'_m.$$

1.2. $\zeta \equiv x_j \in \vec{x}$ $(1 \leq j \leq l)$, so $\xi[\vec{Q}/\vec{x}] \equiv Q_j$. But $P_j \Rightarrow^\circ_\Delta Q_j$ means that there is a standard sequence $P_j \equiv S_0 \rightarrow_\Delta \; ..... \rightarrow_\Delta S_n \equiv Q_j$ $(n \in \mathbb{N})$. Two cases can arise.

1.2.1. $\forall i \leq n$, $S_i \not\equiv \lambda z.S'$. Then the following reduction sequence

$$\sigma : S_0 M'_1...M'_m \rightarrow_\Delta \; ..... \; \rightarrow_\Delta S_n M'_1...M'_m$$

is standard. Since by induction $M_i[\vec{P}/\vec{x}] \rightarrow^\circ_\Delta N_i[\vec{Q}/\vec{x}]$, there is a standard reduction sequence

$$\tau : S_n M'_1...M'_m \rightarrow^\circ_\Delta S_n N'_1 M'_2...M'_m \rightarrow^\circ_\Delta \; ..... \; \rightarrow^\circ_\Delta S_n N'_1...N'_m.$$

Note that $S_0 M'_1...M'_m \equiv M[\vec{P}/\vec{x}]$ and $S_n N'_1...N'_m \equiv N[\vec{Q}/\vec{x}]$, so σ followed by τ is the desired standard reduction sequence.

1.2.2. There is a minimum $k \leq n$ such that $S_k \equiv \lambda z.S'$.
By induction on (ii), $M_1 \Rightarrow^\circ_\Delta N_1$. Therefore, by induction $M_1[\vec{P}/\vec{x}] \Rightarrow^\circ_\Delta N_1[\vec{Q}/\vec{x}]$, where $M_1[\vec{P}/\vec{x}] \rightarrow^\circ_\Delta N_1[\vec{Q}/\vec{x}]$ is $M_1[\vec{P}/\vec{x}] \equiv R_0 \rightarrow_\Delta \; ..... \rightarrow_\Delta R_p \equiv N_1[\vec{Q}/\vec{x}]$ $(p \in \mathbb{N})$. There are two subcases:

1.2.2.1. $\forall i \le p,\ R_i \notin \Delta$. Then the following reduction sequence:

$$\begin{aligned}\sigma' : M[\vec{P}/\vec{x}] \equiv S_0 R_0 M_2' ... M_m' \rightarrow_\Delta ..... \rightarrow_\Delta S_k R_0 M_2' ... M_m' \\ \rightarrow_\Delta ..... \rightarrow_\Delta S_k R_p M_2' ... M_m' \\ \rightarrow_\Delta S_{k+1} R_p M_2' ... M_m' \rightarrow_\Delta ..... \rightarrow_\Delta S_n R_p M_2' ... M_m'\end{aligned}$$

is also standard. Moreover, since $M_i[\vec{P}/\vec{x}] \rightarrow_\Delta^\circ N_i[\vec{P}/\vec{x}]$, the following reduction sequence:

$$\begin{aligned}\tau' : & S_n R_p M_2' ... M_m' \rightarrow_\Delta^\circ \\ & S_n R_p N_2' M_3' ... M_m' \rightarrow_\Delta^\circ \ ..... \ \rightarrow_\Delta^\circ S_n R_p N_2' ... N_m'\end{aligned}$$

is also standard. Clearly σ' followed by τ' is the desired standard reduction sequence.

1.2.2.2. There is a minimum $q \le p$ such that $R_q \in \Delta$. So

$$\begin{aligned}\sigma'' : \quad M[\vec{P}/\vec{x}] \equiv S_0 R_0 M_2' ... M_m' \rightarrow_\Delta ..... \rightarrow_\Delta S_k R_0 M_2' ... M_m' \\ \rightarrow_\Delta ..... \rightarrow_\Delta S_k R_q M_2' ... M_m' \rightarrow_\Delta S_{k+1} R_q M_2' ... M_m' \\ \rightarrow_\Delta ..... \rightarrow_\Delta S_n R_q M_2' ... M_m' \rightarrow_\Delta ..... \rightarrow_\Delta S_n R_p M_2' ... M_m'\end{aligned}$$

is a standard reduction sequence. The desired standard reduction sequence is σ'' followed by τ'.

2. Let $\zeta \equiv (\lambda z.T)U$. Thus $N \equiv (\lambda z.\bar{T})\bar{U} N_1 ... N_m$ or $N \equiv \bar{T}[\bar{U}/z] N_1 ... N_m$, where $T \Rightarrow_\Delta \bar{T}$, $U \Rightarrow_\Delta \bar{U}$ and $M_i \Rightarrow_\Delta N_i$ $(1 \le i \le m)$.

   By induction, $U' \equiv U[\vec{P}/\vec{x}] \Rightarrow_\Delta^\circ \bar{U}[\vec{Q}/\vec{x}] \equiv U''$, $T' \equiv T[\vec{P}/\vec{x}] \Rightarrow_\Delta^\circ \bar{T}[\vec{Q}/\vec{x}] \equiv T''$ and $M_i' \equiv M_i[\vec{P}/\vec{x}] \Rightarrow_\Delta^\circ N_i[\vec{Q}/\vec{x}] \equiv N_i'$ $(1 \le i \le m)$.

   Let $U' \equiv R_0 \rightarrow_\Delta ... \rightarrow_\Delta R_p \equiv U''$ $(p \in \mathbb{N})$ be the standard sequence $U' \rightarrow_\Delta^\circ U''$. Without loss of generality let us assume $z \notin \vec{x}$.

   2.1. Let $N \equiv (\lambda z.\bar{T})\bar{U} N_1 ... N_m$. There are two cases.

   2.1.1. $\forall i \le p \quad R_i \notin \Delta$. Then the standard reduction sequence $M[\vec{P}/\vec{x}] \rightarrow_\Delta^\circ N[\vec{Q}/\vec{x}]$ is

$$\begin{aligned}(\lambda z.T')R_0 M_1' ... M_m' \rightarrow_\Delta ..... \rightarrow_\Delta (\lambda z.T')R_p M_1' ... M_m' \\ \rightarrow_\Delta^\circ (\lambda z.T'')R_p M_1' ... M_m' \rightarrow_\Delta^\circ (\lambda z.T'')R_p N_1' M_2' ... M_m' \\ \rightarrow_\Delta^\circ \ ..... \ \rightarrow_\Delta^\circ (\lambda z.T'')R_p N_1' ... N_m'.\end{aligned}$$

   2.1.2. There is a minimum $q \le p$ such that $R_q \in \Delta$. Thus the desired standard reduction sequence is:

$$\begin{aligned}(\lambda z.T')R_0 M_1' ... M_m' \rightarrow_\Delta ..... \rightarrow_\Delta (\lambda z.T')R_q M_1' ... M_m' \\ \rightarrow_\Delta^\circ (\lambda z.T'')R_q M_1' ... M_m' \rightarrow_\Delta ..... \rightarrow_\Delta (\lambda z.T'')R_p M_1' ... M_m' \\ \rightarrow_\Delta^\circ (\lambda z.T'')R_p N_1' M_2' ... M_m' \rightarrow_\Delta^\circ \ ..... \ \rightarrow_\Delta^\circ (\lambda z.T'')R_p N_1' ... N_m'.\end{aligned}$$

2.2. Let $N \equiv \bar{T}[\bar{U}/z]N_1...N_m$. So, there is a minimum $q \le p$ such that $R_q \in \Delta$; let μ be the standard reduction sequence:

$$M[\vec{P}/\vec{x}] \equiv (\lambda z.T')R_0 M'_1...M'_m \to_\Delta ..... \to_\Delta (\lambda z.T')R_q M'_1...M'_m \\ \to_\Delta T'[R_q/z]M'_1...M'_m.$$

$T \Rightarrow^\circ_\Delta \bar{T}$, by induction on (ii). Furthermore, since $R_q \Rightarrow^\circ_\Delta U''$, it follows by induction that $T[\vec{P}/\vec{x}][R_q/z] \Rightarrow^\circ_\Delta \bar{T}[\vec{Q}/\vec{x}][U''/z]$.
Let $T[\vec{P}/\vec{x}][R_q/z] \equiv T_0 \to_\Delta ..... \to_\Delta T_t \equiv \bar{T}[\vec{Q}/\vec{x}][U''/z]$ be the corresponding standard reduction sequence. Two subcases can arise:

2.2.1. $\forall i \le t$, $T_i \not\equiv \lambda z.S'$. The desired standard reduction sequence is μ followed by:

$$T'[R_p/z]M'_1...M'_m \equiv T[\vec{P}/\vec{x}][R_p/z]M'_1...M'_m \to_\Delta T_1 M'_1...M'_m \\ \to_\Delta ..... \to_\Delta T_t M'_1...M'_m \to^\circ_\Delta \;..... \;\to^\circ_\Delta T_t N'_1...N'_m \equiv [\vec{Q}/\vec{x}]$$

2.2.2. Let $k \le t$ be the minimum index such that $T_k \equiv \lambda y.T'_k$. The construction of the standard reduction sequence depends on the fact that M_2 may or may not become an input value, but, in every case, it can be easily built as in the previous cases.

(ii) The cases $M \equiv x$ and $M \equiv \lambda z.M'$ are easy.

1. Let $M \equiv PQ \Rightarrow_\Delta P'Q' \equiv N$, $P \Rightarrow_\Delta P'$ and $Q \Rightarrow_\Delta Q'$.
By induction, there are standard sequences $P \equiv P_0 \to_\Delta ... \to_\Delta P_p \equiv P'$ and $Q \equiv Q_0 \to_\Delta ... \to_\Delta Q_q \equiv Q'$.
If $\forall i \le p \;\; P_i \not\equiv \lambda z.P'_i$, then $M \to^\circ_\Delta N$ is $P_0Q_0 \to^\circ_\Delta P_pQ_0 \to^\circ_\Delta P_pQ_q$.
Otherwise, let k be the minimum index such that $P_k \equiv \lambda z.P'_k$.
- If $\forall j \le q \;\; Q_j \notin \Delta$, then $M \to^\circ_\Delta N$ is

$$P_0Q_0 \to_\Delta ..... \to_\Delta P_kQ_0 \to^\circ_\Delta P_kQ_q \to_\Delta P_{k+1}Q_q \to_\Delta ..... \to_\Delta P_pQ_q.$$

- If there is a minimum h such that $Q_h \in \Delta$, the standard sequence is $P_0Q_0 \to^\circ_\Delta P_kQ_0 \to^\circ_\Delta P_kQ_h \to_\Delta P_{k+1}Q_h \to^\circ_\Delta P_pQ_h \to^\circ_\Delta P_pQ_q$.

2. Let $M \equiv (\lambda x.P)Q \Rightarrow_\Delta P'[Q'/x] \equiv N$ where $P \Rightarrow_\Delta P'$, $Q \Rightarrow_\Delta Q'$ and $Q \in \Delta$. Hence $P \Rightarrow^\circ_\Delta P'$ and $Q \Rightarrow^\circ_\Delta Q'$ follow by induction, so $P[Q/x] \Rightarrow^\circ_\Delta P'[Q'/x]$, by induction on (i). Thus, the desired standard reduction sequence is $(\lambda x.P)Q \to_\Delta P[Q/x] \to^\circ_\Delta P'[Q'/x]$. □

In order to prove the standardization theorem some auxiliary definitions are necessary.

Definition 1.2.27. *Let $M, N \in \Lambda$.*

(i) $M \to^i_\Delta N$ *denotes that N is obtained from M by reducing a redex that is not the principal redex.*
(ii) $M \Rightarrow^i_\Delta N$ *denotes* $M \Rightarrow_\Delta N$ *and* $M \to^{*i}_\Delta N$.

According to this new terminology, a set of input values is standard, in the sense of Definition 1.2.9 (pag. 10), if and only if $M \notin \Delta$ and $M \rightarrow^{*i}_{\Delta} N$ imply $N \notin \Delta$.

Lemma 1.2.28. *$M \Rightarrow_{\Delta} N$ implies there is P such that $M \rightarrow^{*p}_{\Delta} P \Rightarrow^{i}_{\Delta} N$.*

Proof. Trivial, by Lemma 1.2.26.(ii).
Notice that it can be $M \equiv P$, by definition of $\rightarrow^{*p}_{\Delta}$. □

Example 1.2.29. Let $M \equiv (\lambda xy.I(\lambda z.IK(II)))I \Rightarrow_{\Gamma} \lambda yz.IKI$.
Therefore $M \rightarrow^{p}_{\Gamma} \lambda y.I(\lambda z.IK(II)) \rightarrow^{p}_{\Gamma} \lambda yz.IK(II) \Rightarrow^{i}_{\Gamma} \lambda yz.IKI$ and clearly $\lambda yz.IK(II) \in \Gamma$.

Note that if Δ is standard and R is the principal redex of M and $M \rightarrow^{*i}_{\Delta} N$, then R is the principal redex of N.

Lemma 1.2.30. *Let Δ be standard.*
$M \Rightarrow^{i}_{\Delta} P \rightarrow^{p}_{\Delta} N$ implies $M \rightarrow^{*p}_{\Delta} Q \Rightarrow^{i}_{\Delta} N$, for some Q.

Proof. By induction on M. If either $M \equiv \lambda x.M'$, or the head of M is a variable, then the proof follows by induction. Otherwise, let $M \equiv (\lambda y.M_0)M_1...M_m$; thus it must be $P \equiv (\lambda y.P_0)P_1...P_m$. Note that $M \Rightarrow^{i}_{\Delta} P$ implies $M_i \Rightarrow_{\Delta} P_i$ $(1 \leq i \leq m)$. Now there are two cases, according to whether $P_1 \in \Delta$ or not.

- Let $P_1 \in \Delta$; it follows that P_1 is the argument of the principal redex of P, thus $N \equiv P_0[P_1/y]P_2...P_m$.
  Let $M_1 \in \Delta$. Then we can build the following reduction sequence:
  $M \equiv (\lambda y.M_0)M_1...M_m \rightarrow^{p}_{\Delta} M_0[M_1/y]...M_m \Rightarrow_{\Delta} P_0[P_1/y]P_2...P_m$, which can be transformed into a standard one by Lemma 1.2.28.
  Let $M_1 \notin \Delta$ and $P_1 \in \Delta$; since the set Δ is standard, $M_1 \Rightarrow_{\Delta} P_1 \in \Delta$ if and only if $M_1 \rightarrow^{*p}_{\Delta} P'_1 \Rightarrow^{i}_{\Delta} P_1$, where $P'_1 \in \Delta$. But this would imply that in the reduction $M \Rightarrow^{i}_{\Delta} P$ the principal redex of M_1 has been reduced; but by definition the principal redex of M_1 coincides with the principal redex of M, against the hypothesis that $M \Rightarrow^{i}_{\Delta} P$. So this case is not possible.
- Let $P_1 \notin \Delta$. Then there is $j \geq 0$ such that the principal redex of P_j is the principal redex of P. Let $j \geq 2$; so $\forall k \leq j$ P_k is a normal form. So $N \equiv (\lambda y.P_0)P_1...P'_j..P_m$, where $P_j \rightarrow^{p}_{\Delta} P'_j$. From the hypothesis that $M \Rightarrow^{i}_{\Delta} P$, it follows that $M_i \equiv P_i$ $(0 \leq i \leq j-1)$, and $M_i \Rightarrow_{\Delta} P_i$ $(j < i \leq m)$. Then by induction there is P^*_j such that $M_j \rightarrow^{*p}_{\Delta} P^*_j \Rightarrow^{i}_{\Delta} P'_j$, and we can build the following reduction sequence:

$$(\lambda y.M_0)M_1...M_m \rightarrow^{*p}_{\Delta} (\lambda y.M_0)M_1...P^*_j P_{j+1}...P_m \Rightarrow_{\Delta} (\lambda y.M_0)M_1...P'_j...P_m$$

  which can be transformed into a standard one by Lemma 1.2.28.
  The case $j < 2$ is similar. □

This Lemma has a key corollary.

Corollary 1.2.31. *Let Δ be standard.*
If $M \to_\Delta^* N$ then $M \to_\Delta^{*p} Q \underbrace{\Rightarrow_\Delta^i \dots \Rightarrow_\Delta^i}_{k} N$, for some Q and some k.

Proof. Note that if $P \to_\Delta P'$ then $P \Rightarrow_\Delta P'$. So $M \to_\Delta^* N$ implies $M \Rightarrow_\Delta N_1 \Rightarrow_\Delta ... \Rightarrow_\Delta N_n \Rightarrow_\Delta N$. So, by repeatedly applying Lemma 1.2.28 and Lemma 1.2.30 we reach the proof. □

Now we are able to prove the standardization theorem.

▼ **Proof of Standardization Theorem** (Theorem 1.2.10 pag. 10).
The proof is given by induction on N. From Corollary 1.2.31, $M \to_\Delta^* N$ implies $M \to_\Delta^{*p} Q \to_\Delta^{*i} N$ for some Q. Obviously, the reduction sequence $\sigma : M \to_\Delta^{*p} Q$ is standard by definition of $\to_\Delta^p$. Note that, by definition of $\to_\Delta^{*i}$, $Q \to_\Delta^{*i} N$ implies that Q and N have the same structure, i.e. $Q \equiv \lambda x_1...x_n.\zeta Q_1...Q_n$ and $N \equiv \lambda x_1...x_n.\zeta' N_1...N_n$, where $Q_i \to_\Delta^* N_i$ $(i \leq n)$ and either ζ and ζ' are the same variable, or $\zeta \equiv (\lambda x.R)S$, $\zeta' \equiv (\lambda x.R')S'$, $R \to_\Delta^* R'$ and $S \to_\Delta^* S'$.
The case when ζ is a variable follows by induction. Otherwise, by induction there are standard reduction sequences $\sigma_i : Q_i \to_\Delta^\circ N_i$ $(1 \leq i \leq n)$, $\tau_R : R \to_\Delta^\circ R'$ and $\tau_S : S \to_\Delta^\circ S'$. Let $S \equiv S_0 \to_\Delta ..... \to_\Delta S_k \equiv S'$ $(k \in \mathbb{N})$.
If $\forall i \leq k \;\; S_i \notin \Delta$ then the desired standard reduction sequence is σ followed by $\tau_S, \tau_R, \sigma_1, ..., \sigma_n$.
Otherwise, there is $S_h \in \Delta$ $(h \leq k)$. In this case, let $\tau_S^0 : S_0 \to_\Delta ..... \to_\Delta S_h$ and $\tau_S^1 : S_{h+1} \to_\Delta ..... \to_\Delta S_k$; the desired standard reduction sequence is σ followed by $\tau_S^0, \tau_R, \tau_S^1, \sigma_1, ..., \sigma_n$. ■

1.3 Δ-Theories

In order to model computation, Δ-equality is too weak. As an example, let Δ be either Λ or Γ. If we want to model the termination property, both the terms DD and $(\lambda x.xxx)(\lambda x.xxx)$ represent programs that run forever, while the two terms are $\neq_\Delta$ each other. Indeed $DD \to_\Delta DD$ and $(\lambda x.xxx)(\lambda x.xxx) \to_\Delta (\lambda x.xxx)(\lambda x.xxx)(\lambda x.xxx)$. So it would be natural to consider them equal in this particular setting. But if we want to take into account not only termination but also the size of terms, they need to be different; in fact, the first one reduces to itself while the second increases its size during the reduction. As we will see in the following, for all instances of Δ we will consider, all interesting interpretations of the calculus also equate terms that are not $=_\Delta$.

Let us introduce the notion of Δ-theory.

Definition 1.3.1. (i) *$\mathcal{T} \subseteq \Lambda \times \Lambda$ is a* congruence *whenever:*
- *$(M, M) \in \mathcal{T}$ for each $M \in \Lambda$,*

- $(M, N) \in \mathcal{T}$ *implies* $(N, M) \in \mathcal{T}$,
- $(M, P) \in \mathcal{T}$ *and* $(P, N) \in \mathcal{T}$ *imply* $(M, N) \in \mathcal{T}$,
- $(M, N) \in \mathcal{T}$ *implies* $(C[M], C[N]) \in \mathcal{T}$, *for all contexts* $C[.]$.

(ii) $\mathcal{T} \subseteq \Lambda \times \Lambda$ *is a* Δ-theory *if and only if it is a congruence and* $M =_\Delta N$ *implies* $(M, N) \in \mathcal{T}$.

We will denote $(M, N) \in \mathcal{T}$ also by $M =_\mathcal{T} N$.

Clearly a Δ-theory equating all terms would be completely uninteresting. So we will ask for consistency.

Definition 1.3.2. (i) *A Δ-theory $\mathcal{T}$ is* consistent *if and only if there are $M, N \in \Lambda$ such that $M \neq_\mathcal{T} N$. Otherwise $\mathcal{T}$ is inconsistent.*

(ii) *A Δ-theory $\mathcal{T}$ is* input consistent *if and only if there are $M, N \in \Delta$ such that $M \neq_\mathcal{T} N$. Otherwise $\mathcal{T}$ is input inconsistent.*

(iii) *A Δ-theory $\mathcal{T}$ is* maximal *if and only if it has no consistent extension, i.e. for all $M, N \in \Lambda$ such that $M \neq_\mathcal{T} N$, any Δ-theory $\mathcal{T}'$ containing $\mathcal{T}$ and such that $M =_{\mathcal{T}'} N$ is inconsistent.*

Property 1.3.3. Let $\mathcal{T}$ be a Δ-theory.
If $\mathcal{T}$ is input consistent then it is consistent.

Proof. Obvious. □

In the last section of this book, we will see that in order to use a $\lambda\Delta$-calculus for computing, we need to work inside theories that are both consistent and input consistent.

Δ-theories can be classified according to their behaviour with respect to the Δ-solvable terms.

Definition 1.3.4. (i) *A Δ-theory is* sensible *if it equates all Δ-unsolvable terms.*

(ii) *A Δ-theory is* semisensible *if it never equates a Δ-solvable term and a Δ-unsolvable term.*

Another important notion for Δ-theories is that of separability. In fact, this property help us to understand what equalities cannot be induced by a theory.

Definition 1.3.5. *Let Δ be a set of input values.*
Two terms M, N are Δ-separable *if and only if there is a context $C[.]$ such that $C[M] =_\Delta x$ and $C[N] =_\Delta y$ for two different variables x and y.*

Property 1.3.6. Let M, N be Δ-separable.
If $\mathcal{T}$ is a Δ-theory such that $M =_\mathcal{T} N$ then $\mathcal{T}$ is input inconsistent.

Proof. Let $C[.]$ be the context separating M and N, i.e. $C[M] =_\Delta x$ and $C[N] =_\Delta y$ for two different variables x and y. Since $=_\mathcal{T}$ is a congruence, $M =_\mathcal{T} N$ implies $C[M] =_\mathcal{T} C[N]$, and so, since $\mathcal{T}$ is closed under $=_\Delta$, $x =_\mathcal{T} y$. But this implies $\lambda xy.x =_\mathcal{T} \lambda xy.y$, i.e. $K =_\mathcal{T} O$. But, since $=_\mathcal{T}$ is a congruence, this implies $KMN =_\mathcal{T} OMN$ for all terms M, N. In particular, if $M, N \in \Delta$ then $M =_\mathcal{T} N$ by Δ-reduction. □

A theory is *fully extensional* if all terms in it (not only abstractions) have a functional behaviour. So, in a fully extensional theory, the equality between terms must be extensional (in the usual sense), i.e., it must satisfy the property:

$$\text{(EXT)} \qquad Mx = Nx \Rightarrow M = N \qquad\qquad x \notin \mathrm{FV}(M) \cup \mathrm{FV}(N).$$

Clearly $=_\Delta$ does not satisfy (EXT). In fact, (EXT) holds for $=_\Delta$ only if it is restricted to terms that reduce to an abstraction: indeed, $xy =_\Delta (\lambda z.xz)y$, but $x \neq_\Delta \lambda z.xz$.

The least extensional extension of $=_\Delta$ is induced by the η-reduction rule, defined as follows:

Definition 1.3.7 (η-Reduction).

(i) *The η-reduction ($\rightarrow_\eta$) is the contextual closure of the following rule:* $\lambda x.Mx \rightarrow_\eta M$ *if and only if* $x \notin \mathrm{FV}(M)$*;*
$\lambda x.Mx$ is a η-redex and M is its contractum;

(ii) $M \rightarrow_{\Delta\eta} N$ *if N is obtained from M by reducing either a Δ or a η redex in M;*

(iii) $\rightarrow^*_{\Delta\eta}$ *and* $=_{\Delta\eta}$ *are respectively the reflexive and transitive closure of* $\rightarrow_{\Delta\eta}$ *and the symmetric, reflexive and transitive closure of* $\rightarrow_{\Delta\eta}$.

The next theorem shows an interesting result for η-reduction.

Theorem 1.3.8. $=_{\Delta\eta}$ *is the least extensional extension of* $=_\Delta$.

Proof. It is immediate to check that $=_{\Delta\eta}$ is extensional. In fact, for $x \notin \mathrm{FV}(M)$, $Mx =_{\Delta\eta} Nx$ implies $\lambda x.Mx =_{\Delta\eta} \lambda x.Nx$ (since $=_{\Delta\eta}$ is a congruence), and this implies, $M =_{\Delta\eta} N$ by $=_\eta$.
On the other hand, let $\mathcal{T}$ be a fully extensional Δ-theory, i.e. $Mx =_\mathcal{T} Nx$ implies $M =_\mathcal{T} N$. For $x \notin \mathrm{FV}(M)$, $(\lambda x.Mx)x =_\mathcal{T} Mx$, since $(\lambda x.Mx)x \rightarrow_\Delta Mx$, and thus by (EXT), $\lambda x.Mx =_\mathcal{T} M$. So $\mathcal{T}$ is closed under $=_\eta$. □

In the literature, full extensionality is called simply extensionality. We use this name to stress the fact that it is also possible to define weaker notions of extensionality. We will develop this topic in Sect. 8.1.