Skeleton for the Proof development leading to the Fundamental Theorem of Algebra

Herman Geuvers, Randy Pollack, Freek Wiedijk, Jan Zwanenburg

November 10, 2000

1 Setoids

A basic ingredient of constructive real numbers is the *apartness* relation $\#$. This is a constructive version of the (classical) inequality on reals: two real numbers are apart if it can positively be decided that they are distinct from each other. In constructive analysis, the apartness is more basic then equality. We therefore take the notion of apartness as a basic ingredient of our structures Usually this apartness is taken to be *tight*, saying that the negation of apartness is the equality. In [3] and [4] also apartness relations occur that are not necessarily tight, but in the formalization of reals one can restrict to a tight apartness. This also implies that, in the formalization, we could have done without an equality alltogether (and replace it with the negation of $\#$). For reasons of exposition and for relating to a more classical set-up we choose to take equality as a primitive.

Definition 1.1 *A binary relation $\#$ on a set S is an* apartness *relation if*

1. *$\#$ is* consistent*, i.e. $\neg a \# a$ for all a.*
2. *$\#$ is* symmetric*, i.e. $a \# b \to b \# a$ for all a, b.*
3. *$\#$ is cotransitive, i.e. $a \# b \to \forall z[a \# z \vee z \# b]$ for all a, b.*

An apartness relation is tight *if its negation is the equality, i.e. $\neg(a \# b) \leftrightarrow a = b$ for all a, b.*

Fact 1.2 *The negation of an apartness relation on S is an equivalence relation on S which is stable, i.e. $\neg\neg\neg(a \# b) \to \neg(a \# b)$.*

Lemma 1.3 *A tight apartness relation respects the equality, i.e.*

$$a \# b \wedge b = b' \to a \# b' \text{ for all } a, b, b'.$$

Proof If $a \# b$, then $a \# b' \vee b \# b'$. The latter is false, because $b = b'$. ■

Definition 1.4 *A* constructive setoid *is a triple $\langle S, =, \# \rangle$, with S a set, $=$ an equivalence relation on S and $\#$ a tight apartness relation on S.*

In a structure, we want the operations and relations to respect the equality and the apartness. For the equality this means that we want to have the *replacement property* for all predicates:

$$R(x_1, \ldots, x_n) \wedge x_1 = y_1 \wedge \ldots \wedge x_n = y_n \to R(y_1, \ldots, y_n).$$

Fact 1.5 *The replacement property is closed under $\vee, \wedge, \neg, \to, \exists$ and $\forall$.*

So, we only have to require that the basic relations satisfy the replacement property and that all basic operations are *well-defined* with respect to the equality, i.e. for f of arity n we have the following.

$$x_1 = y_1 \wedge \ldots \wedge x_n = y_n \to f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n).$$

If we have a tight apartness relation, this immediately implies

$$\neg(x_1 \# y_1) \wedge \ldots \wedge \neg(x_n \# y_n) \to \neg(f(x_1, \ldots, x_n) \# f(y_1, \ldots, y_n)),$$

but one would like to have a more positive formulation saying

$$f(x_1, \ldots, x_n) \# f(y_1, \ldots, y_n) \to (x_1 \# y_1) \vee \ldots \vee (x_n \# y_n).$$

This property is called *strong extensionality* of f.

Definition 1.6 *Let S be a set with an apartness relation $\#$ defined on it. For f a n-ary function on S, we say that f is* strongly extensional *if*

$$\forall x_1, \ldots, x_n, y_1, \ldots, y_n[f(\vec{x}) \# f(\vec{y}) \to (x_1 \# y_1 \vee \ldots \vee x_n \# y_n)].$$

For R a n-ary relation on S, we say that R is strongly extensional *if*

$$\forall x_1, \ldots, x_n, y_1, \ldots, y_n[R(\vec{x}) \to (R(\vec{y}) \vee x_1 \# y_1 \vee \ldots \vee x_n \# y_n)].$$

Fact 1.7 *Strong extensionality of functions is closed under composition. Strong extensionality of relations is closed under $\vee$, $\wedge$, $\exists$ and the substitution of strongly extensional terms.*

Lemma 1.8 *Strong extensionality implies well-definedness for functions.*

Proof Suppose $f(\vec{x}) \# f(\vec{y}) \to (x_1 \# y_1 \vee \ldots \vee x_n \# y_n)$ for all $x_1, \ldots, x_n, y_1, \ldots, y_n$. Suppose $x_1 = y_1 \wedge \ldots \wedge x_n = y_n$ and $f(\vec{x}) \# f(\vec{y})$. Then $x_1 \# y_1 \vee \ldots \vee x_n \# y_n$ by strong extensionality of f. Contradiction, so $\neg(f(\vec{x}) \# f(\vec{y}))$, i.e. $f(\vec{x}) = f(\vec{y})$. ■

Remark 1.9 *Strong extensionality (for functions) says that a function can only distinguish elements that can already be distinguished. We will require all basic functions in constructive structures to be strongly extensional. As a consequence, all composed functions will be strongly extensional.*

We do not want all relations to be strongly extensional. For example, equality is not strongly extensional: if it were, then $x = y \to p = q \vee x \# p \vee y \# q$ for all x, y, p, q, which implies the decidability of equality (take x for y and p).

Lemma 1.10 *If a binary function f is strongly extensional in both arguments, i.e.*

$$\forall x_1, x_2, y[f(x_1, y) \# f(x_2, y) \to (x_1 \# x_2)],$$
$$\forall x, y_1, y_2[f(x, y_1) \# f(x, y_2) \to (y_1 \# y_2)],$$

then it is strongly extensional. Similarly for functions of higher arity.

Proof Suppose the binary function f is strongly extensional in both arguments and suppose $f(x_1, y_1) \# f(x_2, y_2)$. Then $f(x_1, y_1) \# f(x_1, y_2) \vee f(x_1, y_2) \# f(x_2, y_2)$ by cotransitivity. Hence $y_1 \# y_2 \vee x_1 \# x_2$. ■

Lemma 1.11 *If f is strongly extensional, then*

$$f(\vec{x}) \neq f(\vec{y}) \to \neg(x_1 = y_1 \wedge \ldots \wedge x_n = y_n).$$

Proof Suppose $f(\vec{x}) \neq f(\vec{y})$, i.e. $\neg\neg(f(\vec{x}) \neq f(\vec{y}))$. Suppose also that $x_1 = y_1 \wedge \ldots \wedge x_n = y_n$. Now, if $f(\vec{x}) \# f(\vec{y})$, then $x_1 \# y_1 \vee \ldots \vee x_n \# y_n$, contradicting $x_1 = y_1 \wedge \ldots \wedge x_n = y_n$. So $\neg(f(\vec{x}) \# f(\vec{y}))$, contradicting $f(\vec{x}) \neq f(\vec{y})$. So we conclude that $\neg(x_1 = y_1 \wedge \ldots \wedge x_n = y_n)$. ■

If a function f has an inverse, we want it to *respect* the apartness. Note that, if f has no inverse we do not want that in general (e.g. consider multiplication in $\mathbb{Z}_4$). That f respects $\#$ comes as a consequence of strong extensionality and the existence of an inverse.

Lemma 1.12 *Suppose that the unary function f has an inverse g which is strongly extensional. Then f respects the apartness, i.e.*

$$x \# y \to f(x) \# f(y).$$

Proof We know that $g(x) \# g(y) \to x \# y$ and that $g(f(x)) = x$. Now suppose $x \# y$, i.e. $g(f(x)) \# g(f(y))$. Then $f(x) \# f(y)$ by strong extensionality of g. ■

Lemma 1.13 *If f respects the apartness, then f respects the inequality*

Proof Let f respect the apartness(i.e. $(x_1 \# y_1 \vee \ldots \vee x_n \# y_n) \to f(\vec{x}) \# f(\vec{y})$). Suppose $x_1 \neq y_1 \vee \ldots \vee x_n \neq y_n$ and suppose $f(\vec{x} = f(\vec{y})$. Now, if $x_i \# y_i$ for some i, then $f(\vec{x} \# f(\vec{y})$, contradiction, so $x_i = y_i$ for all i. This is again a contradiction, so $f(\vec{x}) \neq f(\vec{y})$. ■

Lemma 1.14 *If a relation R is strongly extensional in each of its arguments, it is strongly extensional.*

Proof We give the proof for a binary relation R. Suppose R is strongly extensional in both arguments, i.e.

$$\begin{aligned} R(x,y) &\to R(x,q) \vee y \# q, \\ R(x,y) &\to R(p,y) \vee p \# x. \end{aligned}$$

for all x, y, p, q. Now, if $R(x,y)$, then $R(x,q) \vee y \# q$. If $R(x,q)$, then $R(p,q) \vee p \# x$, so $R(p,q) \vee p \# x \vee y \# q$. ■

Lemma 1.15 *Apartness is strongly extensional.*

Proof Suppose $x \# y$. Then $x \# p \vee y \# p$ by cotransitivity and hence $x \# p \vee y \# q \vee p \# q$ by again cotransitivity. ■

1.1 On the choice of the primitives

In view of the fact that we require an apartness relation in a setoid to be tight, we could have chosen to define a setoid as a pair $\langle S, \# \rangle$ with $\#$ an apartness relation and then *define* equality by

$$x = y \quad := \quad \neg(x \# y).$$

Then the following can be shown.

1. If an operation f is strongly extensional, then it respects $=$.
2. If a relation R is strongly extensional, then it satisfies the replacement property.
3. Hence all relations satisfy the replacement property.

So, we could have done without an equality alltogether. However, we have not chosen this option, because equality is a natural primitive. Furthermore one may at some point encounter structures in which apartness is not tight.

1.2 Subsetoids and Quotient Setoids

Definition 1.16 *Given a constructive setoid $\langle S, =, \# \rangle$ and a predicate P on S, we define the* subsetoid of the $x \in S$ that satisfy P *as the setoid $\langle \{x \in S \mid P(x)\}, =', \#' \rangle$, where $='$ and $\#'$ are the equality and apartness inherited from S, i.e. for $q, t \in \{x \in S \mid P(x)\}$,*

$$\begin{aligned} t =' q &\iff t = q, \\ t \#' q &\iff t \# q, \end{aligned}$$

We denote this subsetoid just by $\{x \in S \mid P(x)\}$.

For this definition to be correct, it has to be shown that $='$ is indeed an equivalence relation and that $\#'$ is a tight apartness relation (w.r.t. $=$) on $\{x \in S \mid P(x)\}$. This is trivially the case. As the equivalence and apartness are directly inherited from S, we never write them explicitly, but use the ones from S.

Definition 1.17 *Given a constructive setoid $\langle S, =, \# \rangle$ and a strongly extensional apartness relation Q on S, we define the* co-quotient setoid S/R *as the setoid $\langle S, \overline{R}, R \rangle$, where $\overline{R}$ is the complement of R, i.e. $\overline{R}(x, y)$ iff $\neg R(x, y)$.*

For this definition to be correct, it has to be shown that $\overline{R}$ is an equivalence relation and that R is a tight apartness relation (w.r.t. $\overline{R}$) on $\{x \in S \mid P(x)\}$. This follows trivially from the definition of $\overline{R}$ and the fact that R is an apartness.

If we do not require R to be strongly extensional, S/R as defined above is still a constructive setoid. However, we only want to consider the situation where the new apartness R is a subset of the old one, i.e. $R(x, y) \to x \# y$. This is a consequence of strong extensionality of R: take x for p and for q in $R(x, y) \to (R(p, q) \vee x \# p \vee y \# q)$. As a consequence we then find that $=$ is a subset of $\overline{R}$, so the new equality is a refinement of the old one. So, the definition of co-quotient setoid subsumes the ordinary definition of quotient set.

For a strongly extensional function f on a setoid $\langle S, =, \# \rangle$ we find that, if f is strongly extensional w.r.t. R, with R a strongly extensional apartness relation on S, then f is also strongly extensional on the co-quotient setoid.

The real numbers form a primary example of a co-quotient setoid. They can be seen as the setoid $(\mathbb{N} \to \mathbb{Q})/R$, where $\mathbb{N} \to \mathbb{Q}$ is the set of infinite sequences of rational numbers and R is the apartness relation between such sequences: for r and s two sequences, $R(r, s)$ iff $\exists k, N \in \mathbb{N} \forall m > N(|r_m - s_m| > \frac{1}{k})$.

2 Constructive Commutative Algebra

We define the notions of commutative monoid, group, ring, integral domain and field in a constructive way. In doing so, we follow [3] and [6], by requiring the basic operations to be strongly extensional. In the end this choice does not effect our work, because in a real number structure, it can be proved from the axioms that all basic operations and relations are strongly extensional.

Convention 2.1 *Without stating it explicitly, we require all operations on setoids to respect the equality. We also require all basic operations and relations on setoids to be strongly extensional (Definition 1.6).*

2.1 Groups: One associative operation

Definition 2.2 (Constructive Semi-Group) *A* constructive semi-group *is a tuple $\langle S, +, =, \# \rangle$ with $\langle S, =, \# \rangle$ a constructive setoid, $+$ a binary operation on S such that*

1. *$+$ is* associative*: $\forall x, y, z[(x + y) + z = x + (y + z)]$.*

Definition 2.3 (Constructive Monoid) *A* constructive monoid *is a tuple* $\langle S, 0, +, =, \#\rangle$ *with* $\langle S, +, =, \#\rangle$ *a constructive semi-group and* 0 *an element of* S *such that*

1. 0 *is the* identity *w.r.t.* $+$*:* $\forall x[x + 0 = x]$.

Definition 2.4 (Constructive Group) *A* constructive group *is a tuple* $\langle S, 0, +, -, =, \#\rangle$ *with* $\langle S, 0, +, =, \#\rangle$ *a constructive monoid,* $-$ *a unary operation on* S *such that*

1. $-x$ *is the inverse of* x*:* $\forall x[x + (-x) = 0]$.

In practice we write "$x - y$*" for "*$x + (-y)$*".*

Lemma 2.5 (Inverses are unique) *The inverse of* $+$ *is unique, i.e. for all* x, y*,*

$$x + y = 0 \to y = -x.$$

As a consequence we find immediately that for all x, y*,*

$$\begin{aligned} -(-x) &= x, \\ -(x + y) &= (-y) + (-x). \end{aligned}$$

Lemma 2.6 (Cancellation) *For all* x, y, z*,*

$$x + y = x + z \to x = y.$$

2.1.1 Apartness in Groups

Lemma 2.7 *For all* x, y*,*

$$x + y \# 0 \to x \# 0 \lor y \# 0.$$

Proof By strong extensionality of $+$, $x + y \# 0 + 0$ implies $x \# 0 \lor y \# 0$. ■

Lemma 2.8 *The operations of a group respect* $\#$*, i.e. for all* x, y, z*,*

$$\begin{aligned} x \# y &\leftrightarrow x + z \# y + z, \\ x \# y &\leftrightarrow x - y \# 0, \\ x \# 0 &\leftrightarrow -x \# 0. \end{aligned}$$

Proof For direction $\to$, $(x + z) - z = x \# y = (y + z) - z$ (using that $\#$ respects $=$, 1.3), so by strong extensionality of $+$, $x + z \# y + z$. The converse uses the forward direction with $-z$. The second part follows from the first part noting that $0 = y - y$. ■

Remark 2.9 *As has already been pointed out, we* always *require functions to respect the equality and to be strongly extensional. In general, you'd want a function to respect the inequality or the apartness only if it has an inverse. See Lemma 1.12.*

2.2 Rings: Two associative operations

Definition 2.10 (Constructive Ring) *A* *non-trivial* constructive ring *is a tuple* $\langle S, 0, 1, +, -, *, =, \#\rangle$ *with* $\langle S, 0, +, -, =, \#\rangle$ *a constructive group and* $\langle S, 1, *, =, \#\rangle$ *a constructive monoid such that*

1. *Non-triviality:* $1 \# 0$.

2. $+$ *distributes over* $*$*:* $\forall x, y, z[x * (y + z) = (x * y) + (x * z)]$.

Notation 2.11 *When dealing with rings we replace the operation* $*$ *by juxtaposition, writing* xy *for* $x * y$.

Lemma 2.12 *For all x, y:*

$$\begin{aligned} x0 &= 0, \\ x(-y) &= -(xy). \end{aligned}$$

Proof The first by cancellation: $x0 = x(0+0) = x0 + x0$. The second (using the first) by uniqueness of inverses 2.5. ■

Lemma 2.13 *For all x, y,*

$$xy \# 0 \quad \rightarrow \quad x \# 0 \wedge y \# 0.$$

Proof Suppose $xy \# 0$. As $*$ is strongly extensional, we know $xy \# x0 \rightarrow y \# 0$ and $xy \# 0y \rightarrow x \# 0$. ■

Definition 2.14 (Constructive Field) *A* constructive field *is a tuple $\langle S, 0, 1, +, -, *, ^{-1}, =, \# \rangle$ such that $\langle S, 0, 1, +, -, *, =, \# \rangle$ is a constructive ring and $^{-1}$ is an operation on the subsetoid $\{x \in S \mid x \# 0\}$, such that*

1. x^{-1} is inverse of x with respect to $*$: $\forall x[x \# 0 \rightarrow xx^{-1} = 1]$.

We have not introduces Integral Domains as a seperate algebraic notion, but constructive fields are indeed constructive integral domains, as they satisfy the following property.

Lemma 2.15 (Integral Domain Property) *In a constructive field we have*

$$\forall x, y[x \# 0 \wedge y \# 0 \rightarrow xy \# 0],$$

that is, a constructive field is a constructive integral domain.

Proof Suppose $x \# 0, y \# 0$. Then $(xy)y^{-1} = x \# 0 = 0y^{-1}$, so $xy \# 0$, by strong extensionality of $*$. ■.

Lemma 2.16 *If $\langle S, 0, 1, +, -, *, ^{-1}, =, \# \rangle$ is a constructive field, then $\langle \{x \in S \,|\, x \# 0\}, 1, *, =, \# \rangle$ forms a constructive monoid.*

Lemma 2.17 *In a constructive field, $*$ respects $\#$, i.e.*

$$\forall x, y, z[x \# y \wedge z \# 0 \rightarrow xz \# yz].$$

Proof Suppose $x \# y$ and $z \# 0$. By lemma 2.8, we have that $x - y \# 0$. Hence $xz - yz = (x - y)z \# 0$, and $zx \# zy$ using lemma 2.8 again. ■

Lemma 2.18 *The following hold in a constructive field.*

$$\begin{aligned} x \neq 0 \wedge y \neq 0 &\rightarrow xy \neq 0, \\ x \neq y \wedge z \neq 0 &\rightarrow xz \neq yz, \\ x \neq 0 \wedge xy = 0 &\rightarrow y = 0. \end{aligned}$$

Proof For the first, suppose $\neg\neg(x \# 0)$ and $\neg\neg(y \# 0)$ and suppose $\neg(xy \# 0)$. If $x \not\# 0$, then if $y \# 0$ we would have $xy \# 0$, contradiction, so $\neg(y \# 0$. But this is a contradiction, so $\neg(x \# 0)$. Contradiction, so $\neg\neg(xy \# 0)$.

For the second, suppose $x \neq y$ and $z \neq 0$. Then $x - y \neq 0$ (using Lemma 1.11). Now, $z(x - y) \neq 0$ using the first and hence $zx \neq zy$ using distributivity and again Lemma 1.11.
For the third, suppose $x \neq 0$ and $xy = 0$. If $y \neq 0$, then $xy \neq 0$ by the first. Contradiction, so $\neg(y \neq 0)$, which implies $y = 0$. ■

Remark 2.19 *It is in general not the case that in a constructive integral domain,*

$$xy = 0 \to x = 0 \vee y = 0.$$

This is just because the $\vee$ has a strong interpretation. A weak counterexample is given by defining the real numbers x and y respectively by the following Cauchy sequences of rationals $(x_i)_{i\in\mathbb{N}}$, resp. $(y_i)_{i\in\mathbb{N}}$. (In this definition we use k_{99} as abbreviation of 'the number k where we have just completed a sequence of 99 9s in the decimal series of π. Similarly $i < k_{99}$ if up to i we have not yet encountered such a sequence of 99 9s.)

$$\begin{array}{rcl}
x_i & := & 2^{-i} \textit{ if } i < k_{99}, \\
x_i & := & 2^{i} \textit{ if } i \geq k_{99} \textit{ and } k_{99} \textit{ is even,} \\
x_i & := & 0 \textit{ if } i \geq k_{99} \textit{ and } k_{99} \textit{ is odd,} \\
y_i & := & 2^{-i} \textit{ if } i < k_99, \\
y_i & := & 2^{i} \textit{ if } i \geq k_{99} \textit{ and } k_{99} \textit{ is odd,} \\
y_i & := & 0 \textit{ if } i \geq k_{99} \textit{ and } k_{99} \textit{ is even.}
\end{array}$$

Now, $xy = 0$, but to say that $x = 0$ implies that we know that k_{99} exists and that it is odd. Similarly for $y = 0$.

Lemma 2.20 *In a constructive field we have the following*

$$\begin{array}{rcl}
x^2 = a^2 & \to & \neg(x \# a \wedge x \# -a), \\
2 \# 0 \wedge a \# 0 \wedge x^2 = a^2 & \to & x = a \vee x = -a.
\end{array}$$

Proof If $x^2 = a^2$, then

$$(x-a)(x+a) = 0. \quad (*)$$

Now, if $x \# a \wedge x \# -a$, then $x - a \# 0 \wedge x + a \# 0$, hence $(x-a)(x+a) \# 0$, contradiction.
If also $a \# 0$, then $a \# -a$ $(a \# 0 \to 2a \# 0 \to a \# -a)$. Hence $x \# a \vee x \# -a$ by cotransitivity of $\#$. Now, if $x \# a$, then $x - a \# 0$ hence $x + a = 0$ by (*) and Lemma 2.18, and hence $x = -a$. Similarly, if $x \# -a$, then $x = a$. ∎

Remark 2.21 *In the previous Lemma we use the premise $2 \# 0$ to conclude $a \not\# -a$ from $a \not\# 0$. Note that if $2 = 0$, the result $(a \# 0 \wedge x^2 = a^2 \to x = a \vee x = -a)$ is also true. We would like to prove $2 \# 0 \vee 2 = 0$, so we can drop the assumption $2 \# 0$ in the Lemma. We conjecture this to hold (the charachteristic of a field is discreet, also constructively (?)). Note that for the present development this doesn't really matter, as we will only be dealing with fields of charachteristic 0 (hence $2 \# 0$): the reals and the complex numbers.*

Lemma 2.22 *If $\langle S, 0, 1, +, -, *, ^{-1}, =, \# \rangle$ is a constructive field, then $\langle \{x \in S \,|\, x \# 0\}, 1, *, ^{-1}, =, \# \rangle$ is a constructive group.*

Definition 2.23 (Constructive Ordered Field) *A* constructive ordered field *is a tuple $\langle S, 0, 1, +, -, *, ^{-1}, =, <, \# \rangle$ such that $\langle S, 0, 1, +, -, *, ^{-1}, =, \# \rangle$ is a constructive field and $<$ is a binary relation on S such that*

1. *$<$ is transitive, irreflexive, anti-symmetric,*
2. *$+$ respects $<$, i.e. $\forall x, y[x < y \to \forall z[x + z < y + z]]$,*
3. *$*$ respects $0 <$, i.e. $\forall x, y[0 < x \wedge 0 < y \to 0 < xy]$,*
4. *$\forall x, y[x \# y \leftrightarrow (x < y \vee y < x)]$,*

2.2.1 Properties of the ordering in a field

Lemma 2.24 (< is cotransitive) *The relation < is cotransitive:*

$$\forall x, y[x < y \to \forall z[x < z \vee z < y]].$$

Proof Suppose $x < y$, then $x \# y$, so for all z: $x \# z \vee z \# y$. Hence $x < z \vee z < x \vee z < y \vee y < z$. As $z < x \to z < y$ and $y < z \to x < z$, we conclude that $x < z \vee z < y$. ■

Lemma 2.25 (∗ respects <)

$$a < b, c > 0 \to ac < bc.$$

Proof Using Axioms 3, 4 and distributivity. ■

Lemma 2.26

$$a > b \leftrightarrow -a < -b.$$

Proof Using Axiom 3. ■

Corollary 2.27

$$0 < 1.$$

Proof $0 < 1 \vee 1 < 0$. If $1 < 0$, then $-1 > 0$ (by 2.26), so $1 = (-1)(-1) > 0$. ■

Lemma 2.28

$$a > 0 \leftrightarrow a^{-1} > 0.$$

Proof If $a > 0$, then $a \# 0$, so $a^{-1} \# 0$, i.e. $a^{-1} > 0 \vee a^{-1} < 0$. Now, if $a^{-1} < 0$, then $-a^{-1} > 0$ (Lemma 2.26), so $-1 > 0$ (Lemma 2.25), contradiction. Hence $a^{-1} > 0$. ■

Lemma 2.29

$$\begin{aligned} a > b > 0 \quad &\leftrightarrow \quad b^{-1} > a^{-1} > 0 \\ 0 > a > b \quad &\leftrightarrow \quad 0 > b^{-1} > a^{-1}. \end{aligned}$$

Proof The first using Lemma 2.25 and Lemma 2.28. The second using Lemma 2.26 and the first. ■

Definition 2.30 *We define the relation ≤ by*

$$x \leq y := \neg(y < x).$$

Lemma 2.31

$$\begin{aligned} x \leq y \quad &\to \quad x + z \leq y + z, \\ x \leq y \wedge z > 0 \quad &\to \quad xz \leq yz. \end{aligned}$$

Proof If $x + z > y + z$, then $x > y$ by Axiom 3. If $xz > yz$ and $z > 0$, then $x > y$ by Lemma 2.25, using Lemma 2.28. ■

Lemma 2.32 $x \geq y \leftrightarrow \forall z[y > z \to x > z]$.

Proof From right to left: Suppose $x < y$, then $x < \frac{x+y}{2} < y$, contradiction.
From left to right: Let $z \in \mathbb{R}$ be such that $y > z$. Then $y > x \vee x > z$. As $y \leq x$, we conclude that $x > z$. ■

Lemma 2.33

$$\begin{aligned} x \leq y \wedge y \leq x \quad &\to \quad x = y, \\ x > y \vee y = x \quad &\to \quad x \geq y \end{aligned}$$

Proof Both trivial. ■

Lemma 2.34

$$\begin{aligned} x < y \wedge y \leq z &\rightarrow x < z, \\ x \leq y \wedge y < z &\rightarrow x < z, \\ x \leq y \wedge y \leq z &\rightarrow x \leq z. \end{aligned}$$

Proof For the first, if $x < y$, then $x < z \vee z < y$. As $y \leq z$ we conclude that $x < z$. The second is similar. For the third, suppose $x > z$. The $x > z$ using the second, contradiction. ■

Lemma 2.35 $x^2 \geq 0$.

Proof Suppose $x^2 < 0$. Then $x^2 \# 0$, so $x \# 0$ (using 2.13), so $x > 0$ or $x < 0$. In the first case $x^2 > 0$. In the second case $-x > 0$, so $(-x)^2 = x^2 > 0$. Contradiction in both cases, so $\neg(x^2 < 0)$. ■

3 The Reals

We give a constructive axiomatization of the reals $\mathbb{R}$. The intention is that the axioms can be instantiated by any specific construction of $\mathbb{R}$. In our axiomatization, the reals form a constructive ordered field, for which Cauchy-completeness and the axiom of Archimedes hold.

Definition 3.1 *A* structure for real numbers *is a constructive abelian ordered field* $\langle \mathbb{R}, 0, 1, +, *, -, ^{-1}, =, <, \# \rangle$ *that*

1. *is Cauchy-complete:*

$$\forall x_1, x_2, \ldots (\forall \epsilon > 0 \exists N \in \mathbb{N} \forall m > N(-\epsilon < x_m - x_N < \epsilon)) \rightarrow \exists x[x = \lim_{n \to \infty} x_n].$$

2. *satisfies the Axiom of Archimedes:*

$$\forall x \exists n \in \mathbb{N}[x < n].$$

Remark 3.2 (to the Definition) *In the definition we actually use* $\mathbb{N}$ *to denote both* $\mathbb{N}$ *itself (in 1) as the* image *of* $\mathbb{N}$ *in* $\mathbb{R}$ *(in 2) via the function* f *that maps* $0 \in \mathbb{N}$ *to* $0 \in \mathbb{R}$ *and* $S(x) \in \mathbb{N}$ *to* $f(x) + 1$*. So, in 1, we quantify over the set of functions from* $\mathbb{N}$ *to* $\mathbb{R}$*. In 2, the axiom really reads* $\forall x \exists n \in \mathbb{N}[x < f(n)]$*, with* f *the injection of* $\mathbb{N}$ *into* $\mathbb{R}$*.*

3.1 Properties of the real numbers

Lemma 3.3 *There are no non-standard real numbers, i.e.*

$$\forall n \in \mathbb{N}[-\frac{1}{n} < x < \frac{1}{n}] \rightarrow x = 0,$$

for all x*.*

Proof Suppose $x \# 0$ towards a contradiction. (Then $\neg(x \# 0)$ and hence $x = 0$.) Then x^{-1} exists and by the Axiom of Archimedes we find a $n \in \mathbb{N}$ such that $\frac{1}{x} < n$. But then either $\frac{1}{n} < x$ or $x < \frac{-1}{n}$ (distinguishing cases according to $x > 0$ or $x < 0$ and using Lemma 2.29. ■

Definition 3.4 *A sequence of reals* $x_1, x_2, \ldots$ *is called a* Cauchy sequence *if*

$$\forall \epsilon > 0 \exists N \in \mathbb{N} \forall m > N(-\epsilon < x_m - x_N < \epsilon).$$

Lemma 3.5 *A sequence of reals $x_1, x_2, \ldots$ is a Cauchy-sequence iff*

$$\forall k \in \mathbb{N} \exists N \in \mathbb{N} \forall m > N(-\frac{1}{k} < x_m - x_N < \frac{1}{k}).$$

Proof The implication from left to right is immediate, as $\frac{1}{k} > 0$. The reverse implication uses the axiom of Archimedes. Assume $\forall k \in \mathbb{N} \exists N \in \mathbb{N} \forall m > N(-\frac{1}{k} < x_m - x_N < \frac{1}{k})$. Let $\epsilon > 0$. Then $\epsilon^{-1} \in \mathbb{R}$, so there is a $k \in \mathbb{N}$ such that $\epsilon^{-1} < k$ and hence $\epsilon > \frac{1}{k}$. Now we find N by our assumption. ∎

Lemma 3.6 *Given $x, y \in \mathbb{R}$ and $\epsilon > 0$, there exists $x' \in \mathbb{R}$ such that*

$$-\epsilon < x - x' < \epsilon \wedge x' \# y.$$

Proof $y < x + \frac{\epsilon}{2} \vee y > x - \frac{\epsilon}{2}$. In the first case take $x' := x + \frac{\epsilon}{2}$, in the second case take $x' := x - \frac{\epsilon}{2}$. ∎

We now define the *maximum* of two real numbers. This is not straightfoward, because we have no trichotomy. (Classically, the maximum can be defined in an ordered field, but constructively that is in general not the case: one needs the Cauchy property.) In a situation where the reals are constructed out of the rationals, say, $x = (x_i)_{i\in\mathbb{N}}$, one can use the maximum of two rationals $(\max(x_i, y_i))$ to define a Cauchy sequence of the maximum of x and y, namely $(\max(x_i, y_i)_{i\in\mathbb{N}}$. Here we can not do that. Instead when defining the maximum of x and y we first have to define an auxiliary sequence of reals $(y_i)_{i\in\mathbb{N}}$ that has y as a limit and such that $x \# y_i$ for all i.

Definition 3.7 *We construct a sequence $(y_i)_{i\in\mathbb{N}}$ such that*

$$\forall i \in \mathbb{N}[-\frac{1}{i} < y - y_i < \frac{1}{i} \wedge y_i \# x].$$

This is is possible, due to Lemma 3.6. Note that $(y_i)_{i\in\mathbb{N}}$ is a Cauchy sequence and $y = \lim_{i\to\infty} y_i$. Now define the sequence $(s_i)_{i\in\mathbb{N}}$ by

$$s_i := \begin{cases} x & \text{if } x > y_i, \\ y_i & \text{if } x < y_i, \end{cases}$$

Now $(s_i)_{i\in\mathbb{N}}$ is a Cauchy sequence and we define

$$max(x, y) := \lim_{i\to\infty} s_i.$$

Lemma 3.8 $\forall x, y \in \mathbb{R}[\neg(max(x, y) > x \wedge max(x, y) > y)]$.

Proof From the Definition of max. ∎

Lemma 3.9 *max is commutative, i.e.* $\forall x, y \in \mathbb{R}[max(x, y) = max(y, x)]$.

Lemma 3.10 *max gives an upperbound, i.e.* $\forall x, y \in \mathbb{R}[max(x, y) \geq x \wedge max(x, y) \geq y]$.

Lemma 3.11 *max give a least upperbound, i.e.* $\forall x, y, z \in \mathbb{R}[z \geq x \wedge z \geq y \to z \geq max(x, y)]$.

Proof Suppose $z < \max(x, y)$. Then $z < y \vee y < \max(x, y)$ and $z < x \vee x < \max(x, y)$. If $y < \max(x, y)$, then $x < \max(x, y)$ contradicts Lemma 3.8, so $z < y \vee z < x$, contradicting $z \geq x \wedge z \geq y$. So $z \geq \max(x, y)$. ∎

Lemma 3.12 $\forall x, y \in \mathbb{R}[x \leq y \leftrightarrow max(x, y) = y]$.

Proof For $\to$: $\max(x, y) \geq y$ by Lemma 3.10 and $\max(x, y) \leq y$ by Lemma 3.11.
For $\leftarrow$: suppose $x > y$. Then $x \geq y$ and hence $\max(x, y) = x$ by the previous. Hence $x = y$, contradiction. So $x \leq y$. ∎

Definition 3.13 *For $x \in \mathbb{R}$, we define*

$$|x| := max(x, -x).$$

Lemma 3.14 $\forall x \in \mathbb{R}[x \geq 0 \to |x| = x]$.

Proof If $x \geq 0$, then $-x \leq 0$, hence $-x \leq x$. So $\max(x, -x) = x$ by Lemma 3.12. ■

Lemma 3.15 $\forall x, y, r \in \mathbb{R}[|x - y| \leq r \leftrightarrow x - r \leq y \leq x + r]$.

Proof Immediate using the intermediate equivalent statement $x - y \leq r \wedge -x + y \leq r$. ■

Lemma 3.16 $\forall x, y \in \mathbb{R}[|x + y| \leq |x| + |y|]$.

Proof $\max(x, -x) + \max(y, -y) \geq x + y$ and $\max(x, -x) + \max(y, -y) \geq -x - y$. Hence $|x| + |y| = \max(x, -x) + \max(y, -y) \geq \max(x + y, -x - y) = |x + y|$. ■

Lemma 3.17 $\forall x, y, z, r, q \in \mathbb{R}[|x - y| \leq r \wedge |y - z| \leq q \to |x - z| \leq r + q$.

Proof $|x - z| = |x - y + y - z| \leq |x - y| + |y - z| \leq r + q$. ■

Lemma 3.18 *If $x_0, x_1, \ldots$ and $y_0, y_1, \ldots$ are Cauchy sequences with limits (respectively) x and y, then*

$$\begin{array}{rcl}
\lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n &=& \lim_{n\to\infty} (x_n + y_n), \\
(\lim_{n\to\infty} x_n)(\lim_{n\to\infty} y_n) &=& \lim_{n\to\infty} (x_n y_n), \\
|\lim_{n\to\infty} x_n| &=& \lim_{n\to\infty} |x_n|, \\
-\lim_{n\to\infty} x_n &=& \lim_{n\to\infty} (-x_n), \\
(\lim_{n\to\infty} x_n)^{-1} &=& \lim_{n\to\infty} (x_n)^{-1},
\end{array}$$

where the latter is only defined if $\forall i \in \mathbb{N}[x_i \# 0]$. Furthermore, if $x_i \leq y_i$ for all i, then $x \leq y$.

In the following, for $n \in \mathbb{N}$, y^n denotes the n-times multiplication of y.

Definition 3.19 *For $x \geq 0$ and $n \in \mathbb{N}^+$, we define $\sqrt[n]{x}$. First we notice that $(x+1)^n > x$ (proof by induction on n). Define the sequences $(p_i)_{i\in\mathbb{N}}$ and $(q_i)_{i\in\mathbb{N}}$ as follows.*

$$\begin{array}{rcl}
p_0 &:=& 0, \\
q_0 &:=& x + 1, \\
p_{i+1} &:=& \begin{cases} p_i & \textit{if} \quad (\frac{2p_i+q_i}{3})^n < x, \\ \frac{2p_i+q_i}{3} & \textit{if} \quad (\frac{p_i+2q_i}{3})^n > x \end{cases} \\
q_{i+1} &:=& \begin{cases} q_i & \textit{if} \quad (\frac{p_i+2q_i}{3})^n > x, \\ \frac{p_i+2q_i}{3} & \textit{if} \quad (\frac{2p_i+q_i}{3})^n < x \end{cases}
\end{array}$$

Then we have the following.

1. $\forall i \in \mathbb{N}[p_i < x < q_i]$,
2. $\forall i \in \mathbb{N}[q_{i+1} - p_{i+1} = \frac{2}{3}(q_i - p_i)]$.

So, $(q_i)_{i\in\mathbb{N}}$ is a Cauchy sequence and we define

$$\sqrt[n]{x} := \lim_{i\to\infty} q_i.$$

Lemma 3.20 $\forall x \in \mathbb{R} \forall n \in \mathbb{N}^+[x \geq 0 \to \sqrt[n]{x} \geq 0]$.

Lemma 3.21 $\forall x, y \in \mathbb{R} \forall n \in \mathbb{N}^+ [x, y \geq 0 \to \sqrt[n]{x}\sqrt[n]{y} = \sqrt[n]{xy}]$.

Lemma 3.22

$$\forall x \forall n \in \mathbb{N}^+ [x \geq 0 \to (\sqrt[n]{x})^n = x].$$

Lemma 3.23 *If $x \geq 0$, then there is a unique $y \geq 0$ such that $y^2 = x$.*

Proof Suppose we have y and z ($y, z \geq 0$) such that $y^2 = x = z^2$ and suppose $y \# z$. Using Lemma ?? we conclude that $y = z \vee y = -z$, hence $y = -z$. Also $y \# 0 \vee z \# 0$, so $(y > 0 \wedge z < 0) \vee (y < 0 \wedge z > 0)$. Contradiction. So $y = z$. ■

4 Polynomials

4.1 Definition and general properties

Definition 4.1 *For R a ring, we define the set of* polynomials over R, $R[X]$, *as the finite lists of elements of R. We define the operations $+$, $*$ and $-$ on $R[X]$. Let $f = \langle f_0, \ldots, f_n \rangle$ and $g = \langle g_0, \ldots, g_m \rangle$ be two polynomials Then $f+g$ and $f-g$ are polynomials of length $max\{m+1, n+1\}$ and $f * g$ is a polynomial of length $m+n+1$ defined as follows.*

$$\begin{aligned}
(f+g)_i &:= f_i + g_i \text{ for } i \leq max\{m, n\}, \\
(f-g)_i &:= f_i - g_i \text{ for } i \leq max\{m, n\}, \\
(f*g)_i &:= \Sigma_{j=0}^{i} f_j * g(i-j) \text{ for } i \leq m+n,
\end{aligned}$$

where it is understood that we take f_j (resp. g_j) to be 0 if $j > n$ (resp. $j > m$). The zero and unit are defined by

$$\begin{aligned}
0 &:= \langle\ \rangle, (\text{the empty sequence}), \\
1 &:= \langle 1 \rangle.
\end{aligned}$$

The apartness relation on $R[X]$ is defined by

$$f \# g := \exists i (f_i \# g_i).$$

Note that we use the terminology *length of a polynomial* when talking about the length of the list of coefficients. The length of a polynomial may not be the same as its *degree* (defined precisely in 4.2)may be 0.

It is easy to see that

$$f = g \leftrightarrow \forall i (f_i = g_i).$$

Definition 4.2 *Let $f(X) = f_n X^n + f_{n-1} X^{n-1} + \ldots + f_1 X + f_0$ be a polynomial.*

1. $f(X)$ has degree k, *notation* $\deg(f) = k$, *if $k \leq n$, $f_k \# 0$ and $i = 0$ for all i with $k < i \leq n$,*

2. $f(X)$ has degree at most k, *notation* $\deg(f) \leq k$, *if $k \leq n$, and $i = 0$ for all i with $k < i \leq n$,*

3. $f(X)$ has degree at least k, *notation* $\deg(f) \geq k$, *if $k \leq n$, $f_k \# 0$.*

So, not all polynomials have a degree: 'degree' is not a function on polynomials but a relation between polynomials and natural numbers. However, it is always the case that the degree of $f(X) = f_n X^n + f_{n-1} X^{n-1} + \ldots + f_1 X + f_0$ is at most *n, and if we know that $f_k \# 0$, it is* at least *k.*

Definition 4.3 *A polynomial $f(X) = f_n X^n + f_{n-1} X^{n-1} + \ldots + f_1 X + f_0$ is called* regular *if for its leading coefficient one has $f_n \# 0$. (That is: the polynomial has a degree, which is the same as its length, n.)*

Lemma 4.4 *For R a ring, $R[X]$ is a ring.*

In the following, unless stated otherwise, R is a ring.

Notation 4.5 *A polynomial $f = \langle f_0, \ldots, f_n \rangle$ will often be denoted by $f(X) = f_n X^n + f_{n-1} X^{n-1} + \ldots + f_0$ or by $f(X) = \Sigma_{j=0}^{n} f_j X^j$.*
The multiplication operation $*$ will usually be omitted.

Definition 4.6 *For every polynomial $f(X) = f_n X^n + f_{n-1} X^{n-1} + \ldots + f_0$ over R we define a function $\overline{f} : R \to R$ in the canonical way:*

$$f(a) := f_n a^n + f_{n-1} a^{n-1} + \ldots + f_0.$$

In the following, we will often just write f for this function $\overline{f}$.

The following two Lemmas already hold for integral domains (rings with the additional property $x \mathrel{\#} 0 \wedge y \mathrel{\#} 0 \to xy \mathrel{\#} 0$, see Lemma 2.15), but we have not introduced that notion here.

Lemma 4.7 *Let F be a field and let $f = f_n X^n + \ldots + f_0$ and $g = g_m X^m + \ldots + g_0$ be polynomials over F. Write $h_{m+n} X^{m+n} + \ldots + h_0$ for fg. Then*

$$f_i g_j \mathrel{\#} 0 \to \exists k[i + j \leq k \leq n + m \wedge h_k \mathrel{\#} 0].$$

Proof See [6], p. 417.

Theorem 4.8 *If F is a field, then $F[X]$ satisfies the integral domain property, i.e. for all $f, g \in F[X]$, if $f \mathrel{\#} 0$ and $g \mathrel{\#} 0$, then $fg \mathrel{\#} 0$.*

Proof Suppose $f = f_n X^n + \ldots + f_0 \mathrel{\#} 0$ and $g = g_m X^m + \ldots + g_0 \mathrel{\#} 0$ and let $h_{m+n} X^{m+n} + \ldots + h_0$ be fg. Then $f_i g_j \mathrel{\#} 0$ for some i, j, but then $h_k \mathrel{\#} 0$ for some k. ■

4.2 Factorization and zeros

Let R be a constructive ring.

Lemma 4.9 *Let $f = f_n X^n + \ldots + f_0$ and $g = g_m X^m + \ldots + g_0$ be polynomials over R. Then there exist $k \in \mathbb{N}, q, r \in R[X]$ such that*

$$(g_m)^k f(X) = q(X) g(X) + r(X)$$

and $r(X)$ has length less then m or 0.

Proof See [6], p. 418.

Theorem 4.10 *Let $f(X) \in R[X]$ and $a \in R$. Then*

$$\exists! q(X) \in R[X](f(X) = (X - a) q(X) + f(a)).$$

Proof By Lemma 4.9, $f(X) = q(X)(X - a) + c$, for some polynomial $q(X)$ and $c \in R$. By taking the value of the function f in a, we find that $c = f(a)$. Furthermore, $q(X) = q_{n-1} X^{n-1} + \ldots + q_0$ and we can determine the coefficients of $q(X)$ uniquely from the equation $f(X) = q(X)(X - a) + f(a)$. ■

Corollary 4.11 *For $f(X) \in R[X]$ and $a \in R$,*

$$(X - a) | f(X) \leftrightarrow f(a) = 0.$$

Moreover, if $f(X)$ has length n and f has $n + 1$ zeros, then $f = 0$.

We now prove that if the polynomial f has degree at least k ($n \geq k > 0$) and we are given $n+1$ distinct elements $(a_i)_{1\leq i\leq n+1}$, then $f(a_i) \# 0$ for one of the i. This will be used to prove the Intermediate Value Theorem for polynomials.

Lemma 4.12 *Let $f(X), g(X) \in R[X]$, both of length n. Let $(a_i)_{0\leq i\leq n-1}$ be distinct elements of R (I.e. $a_i \# a_j$ if $i \neq j$.). If $f(a_i) = g(a_i)$ for all i ($0 \leq i \leq n-1$), then $f = g$.*

Proof The polynomial $h := f - g$ has length n and has n zeros, so $h = 0$ by Corollary 4.11. Hence, $f = g$. ■

Let F be a constructive field.

Lemma 4.13 *Let $f(X) \in R[X]$ of length n and let $(a_i)_{1\leq i\leq n}$ be distinct elements of R. Then*

$$\begin{aligned} f(X) \quad = \quad & f(a_1)\frac{(X-a_2)(X-a_3)\cdots(X-a_n)}{(a_1-a_2)(a_1-a_3)\cdots(a_1-a_n)} + \\ & f(a_2)\frac{(X-a_1)(X-a_3)\cdots(X-a_n)}{(a_2-a_1)(a_2-a_3)\cdots(a_2-a_n)} + \\ & \cdots \\ & f(a_n)\frac{(X-a_1)(X-a_2)\cdots(X-a_{n-1})}{(a_n-a_1)(a_n-a_2)\cdots(a_n-a_{n-1})}. \end{aligned}$$

Proof The right hand side of the equation is a polynomial $h(X)$ of length n (note that all the fractions are defined, because all a_i are distinct). Furthermore f and h agree on all a_i, hence $f = h$ by Lemma 4.12. ■

Lemma 4.14 *Let $f(X) \in R[X]$ of degree at least k ($n \geq k > 0$) and let $(a_i)_{1\leq i\leq n+1}$ be distinct elements of R. Then*

$$f(a_i) \# 0$$

for some i.

Proof Write $f(X) = f_n X^n + \ldots + f_0$. By Lemma 4.13 we find that for the coefficient f_k we have

$$f_k = \Sigma_{1\leq i\leq n+1} f(a_i) h_k^i,$$

where h_k^i is the k-th coefficient of the i-th polynomial as above:

$$\frac{(X-a_1)\cdots(X-a_{i-1})(X-a_{i+1})\cdots(X-a_n)}{(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_{n+1})}$$

As $f_k \# 0$, we find that $f(a_i) \# 0$ for at least one i. ■

4.3 Operations on polynomials

We need some formal operations on polynomials. Let F be an ordered field (to make sure that always $n! \# 0$; as ordered field are infinite this is the case).

Definition 4.15 *For $f(X) = a_n X^n + a_{n-1}X^{n-1} + \ldots + a_1 X + a_0$ a polynomial, we define the* derivative *of f, f' as follows.*

$$f'(X) := na_n X^{n-1} + (n-1)a_{n-1}X^{n-2} + \ldots + a_1.$$

Taking k-times the deriavtive of f is denoted as $f^{(k)}$.

Definition 4.16 *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$ be a polynomial over R and $c \in R$. Define the polynomials $f\tilde{}$ and f_c as follows.*

$$\begin{aligned} f\tilde{}(X) &= a_0 X^n + \ldots + a_{n-1}X + a_n, \\ f_c(X) &= \frac{f^{(n)}(c)}{n!}X^n + \frac{f^{(n-1)}(c)}{(n-1)!}X^{n-1} + \ldots + \frac{f'(c)}{1!}X + f(c) \end{aligned}$$

Lemma 4.17 *Let $f(X) = a_n X^n + a_{n-1}X^{n-1} + \ldots + a_0$ be a polynomial over R and $c \in R$. For the function $\overline{f}$ associated to this polynomial (4.6) we have*

$$\begin{aligned} \overline{f\tilde{}}(x) = x^n \overline{f}(x^{-1}),\ \textit{if } x \# 0, \\ \overline{f\tilde{}}(0) = a_n, \\ \overline{f_c}(x) = \overline{f}(x+c). \end{aligned}$$

Proof We informally write f where we refer to the function $\overline{f}$ etcetera.

$$\begin{aligned} x^n f(\frac{1}{x}) &= x^n(a_n x^{-n} + \ldots + a_0) \\ &= a_n + a_{n-1} + \ldots + a_0 x^n \\ &= f\tilde{}(x). \end{aligned}$$

Clearly $f_c(x)$ is a polynomial of maximal degree n. Hence,

$$f_c(x) = b_n x^n + \ldots + b_0.$$

It follows that

$$\begin{aligned} f(c) &= f_c(0) = b_0, \text{ hence } b_0 = f(c), \\ f'(c) &= f_c'(0) = b_1, \text{ hence } b_1 = f'(c) \\ f^{(2)}(c) &= f_c^{(2)}(0) = 2!b_2, \text{ hence } b_2 = \frac{f^{(2)}(c)}{2!} \\ &\ldots \\ f^{(n)}(c) &= f_c^{(n)}(0) = n!b_n, \text{ hence } b_n = \frac{f^{(n)}(c)}{n!}. \blacksquare \end{aligned}$$

5 Real valued functions

In the proof of the Fundamental Theorem of Algebra, we use a strong version of the Intermediate Value Theorem (with a strong conclusion and a strong premise). This is Theorem 6.1.5 of [5].

Definition 5.1 *Intervals (closed and open) in $\mathbb{R}$.*

Definition 5.2 *Continuity of functions $f : \mathbb{R}^n \to \mathbb{R}$ ($n \in \mathbb{N}$).*

Lemma 5.3 *The identity and a constant function are continuous. Continuity is preserved under $+$, $*$, composition and maximum (of finitely many functions).*

Corollary 5.4 *If $f(X)$ is a polynomial over $\mathbb{R}$, then the associated function f is continuous.*

Theorem 5.5 (Intermediate Value Theorem) *Let $a, b \in \mathbb{R}$, $a < b$, and let f be continuous on $[a, b]$ with $f(a) < 0 < f(b)$. Moreover assume that*

$$\forall x, y \in [a, b](x < y \to \exists z \in [x, y](f(z) \# 0)).$$

Then $\exists z \in [a, b](f(z) = 0)$.

Proof See [5], p. 294.

Corollary 5.6 (Intermediate Value Theorem for regular polynomials) *Let f be a regular polynomial over $\mathbb{R}$ and let $a, b \in \mathbb{R}$ such that $a < b$.*

$$\textit{If } f(a) < 0 \textit{ and } f(b) > 0 \textit{ then } \exists z \in [a,b](f(z) = 0).$$

Proof The premise in Theorem 5.5 is satisfied: if n is the degree of f, we choose $n+1$ distinct points in the interval $[x, y]$; due to Lemma 4.14 the value of f is apart from 0 for one of these points. ■

Proposition 5.7 (Roots of polynomials over $\mathbb{R}$ of odd degree) *Every polynomial of odd degree over $\mathbb{R}$ has a root.*

Proof Let f be a polynomial of odd degree. We only have to show that for x sufficiently small, $f(x) < 0$ and for x sufficiently large, $f(x) > 0$. Then Corollary 5.6 does the job. ■.

Lemma 5.8 (Intermediate Value Theorem for stricly monotonic functions) *If $f : \mathbb{R} \to \mathbb{R}$ is strictly monotonic and continuous on some interval I, and $a, b \in I$ with $a < b$, $f(a) < 0$, $f(b) > 0$, then there is a $c \in (a, b)$ with $f(c) = 0$.*

Proof We show that the premise of Theorem 5.5 is satisfied. Let $x, y \in [a, b]$ with $x < y$. Take $z_1, z_2 \in [x, y]$ such that $z_1 < z_2$. Then $f(z_1) < f(z_2)$ due to the strict monotonicity of f. Hence $f(z_1) < 0 \vee f(z_2) > 0$. ■

Definition 5.9 *Let $n \in \mathbb{N}$, $n \geq 1$, $a_1, \ldots, a_n \in \{x \in \mathbb{R} | x \geq 0\}$ with $a_n = 1$. Define $m : [0, \infty) \to [0, \infty)$ as follows.*

$$m(s) := max\{a_i s^i \mid 1 \leq i \leq n\}.$$

Lemma 5.10 *The function m is strictly monotonic on $(0, \infty)$.*

Proof We prove that for every $x, y \in (0, \infty)$, if $x < y$ then $\frac{m(y)}{m(x)} \geq \frac{y}{x}$. (Then $\frac{m(y)}{m(x)} > 1$ and hence $m(y) > m(x)$.) Let $x, y \in (0, \infty)$, $x < y$, and suppose $\frac{m(y)}{m(x)} < \frac{y}{x}$. Then $\frac{x}{y} m(y) < m(x)$. We conclude that $\forall j \in \{1, \ldots, n\}(\frac{x}{y} a_j y^j < m(x))$ and hence

$$\forall j \in \{1, \ldots, n\}(a_j x^j \neq m(x)).$$

(If $a_j x^j = m(x)$, then $\frac{x}{y} a_j y^j < a_j x^j$ and $y^{j-1} < x^{j-1}$, contradiction.) Also

$$\forall j \in \{1, \ldots, n\}(a_j x^j \leq m(x)).$$

From these two, we conclude that

$$\forall j \in \{1, \ldots, n\}(\neg\neg(a_j x^j < m(x))).$$

From Lemma 3.8 we conclude that

$$\neg \forall j \in \{1, \ldots, n\}(a_j x^j < m(x)).$$

From these two statements we derive a contradiction. (Here we use that the universal quantifier ranges over a finite set.) Viz. Suppose $a_1 x < m(x), a_2 x^2 < m(x), \ldots, a_n x^n < m(x)$. Then $\perp$ using the second. Hence $\neg(a_1 x < m(x))$, which contradicts the first, hence $\neg(a_2 x^2 < m(x))$, which contradicts the first, etcetera until we derive $\neg(a_n x^n < m(x))$ and a contradiction. ■

6 Complex numbers

Definition 6.1 *A structure for the complex numbers consists of*

$$\mathbb{C} = \mathbb{R} \times \mathbb{R},$$

where $\mathbb{R}$ is a structure for the real numbers. On a structure of the complex numbers one defines

$$\begin{array}{rcl} (r,s) +_{\mathbb{C}} (r',s') & = & (r+r', s+s'); \\ (r,s) *_{\mathbb{C}} (r',s') & = & (r*r' - s*s', r*s' + r'*s); \\ 0_{\mathbb{C}} & = & (0,0); \\ 1_{\mathbb{C}} & = & (1,0); \\ i & = & (0,1), \\ (r,s) =_{\mathbb{C}} (r',s') & = & r = r' \wedge s = s', \\ (r,s) \#_{\mathbb{C}} (r',s') & = & r \# r' \vee s \# s'. \end{array}$$

Here $+, *, 0, 1, =, \#$ denote the usual operations and relations on the structure for the reals.

Notation 6.2 *As a corollary of the definition, an element $z = (r,s) \in \mathbb{C}$ will also be denoted by $r + is$.*

Proposition 6.3 1. *With the definitions*

$$\begin{array}{rcl} -_{\mathbb{C}}(r,s) & = & (-r,-s); \\ (r,s)^{-1} & = & (\frac{r}{r^2+s^2}, \frac{-s}{r^2+s^2}). \end{array}$$

a structure for the reals becomes a constructive field.

2. *Moreover $i^2 = -1$.*

Let in the following $\mathbb{C}$ be a structure for the complex numbers.

Definition 6.4 *For $z = (r,s) \in \mathbb{C}$ define*

$$\begin{array}{rcl} |z| & = & \sqrt{(r^2+s^2)}; \\ \overline{z} & = & (r,-s). \end{array}$$

Lemma 6.5 *For $z \in \mathbb{C}$ one has*

$$z \# 0 \leftrightarrow |z| \# 0,$$

where the second $\#$ and 0 are taken from the structure of reals.

Proposition 6.6 *For $z_1, z_2 \in \mathbb{C}$ one has*

$$\begin{array}{rcl} \overline{z_1 *_{\mathbb{C}} z_2} & = & \overline{z_1} *_{\mathbb{C}} \overline{z_2}. \\ |z_1 *_{\mathbb{C}} z_2| & = & |z_1| * |z_2|. \\ |z_1 +_{\mathbb{C}} z_2| & \leq & |z_1| + |z_2|. \end{array}$$

Lemma 6.7 *For $z_1, z_2 \in \mathbb{C}$ one has*

$$\frac{z_1}{z_2} < 0 \rightarrow |z_1 + z_2| = |z_1| - |z_2|,$$

where the statement $\frac{z_1}{z_2} < 0$ denotes that the complex number $\frac{z_1}{z_2}$ is of the form $(r,0)$ with $r < 0$..

Proof $|a+b| = |a(1+\frac{b}{a})| = |a||1+\frac{b}{a}| = |a|(1-\frac{b}{a}) = |a| - |b|$. ∎

Lemma 6.8 *For $z \in \mathbb{C}$ one has*

$$z * \overline{z} = |z|^2.$$

6.1 Roots of complex numbers

We show that for $c \in \mathbb{C}$, $c \# 0$ and $n \in \mathbb{N}^+$, $\sqrt[n]{c}$ exists in $\mathbb{C}$. To prove this we rely just on the following two facts: 1. Every positive number in $\mathbb{R}$ has a square root, 2. Every polynomial of odd degree over $\mathbb{R}$ has a root in $\mathbb{R}$. The proof avoids the use of polar coordinates, exponentials and arctan. We have learned this proof from [2]; thanks to R. Kortram who made us aware of this proof.

Let in the following $\mathbb{C}$ be a structure for the complex numbers.

Lemma 6.9 *For each $c = a + ib \in \mathbb{C}$ with $c \# 0$, there exists a solution to $z^2 = c$. In particular, a solution is given by:*

$$z = \sqrt{\frac{\sqrt{a^2+b^2}+a}{2}} + i\sqrt{\frac{\sqrt{a^2+b^2}-a}{2}} \qquad \textit{for} \quad b \geq 0$$
$$z = \sqrt{\frac{\sqrt{a^2+b^2}+a}{2}} - i\sqrt{\frac{\sqrt{a^2+b^2}-a}{2}} \qquad \textit{for} \quad b \leq 0$$

Proof The second statement, including the fact that all square roots that occur take positive numbers, is a straightforward computation (using that $\sqrt{b^2} = b$ when $b \geq 0$, and that $\sqrt{b^2} = -b$ when $b \leq 0$.) For the first statement, because $c \# 0$, we have either $a \# 0$ or $b \# 0$. The second case we explicitly solved, and the first case reduces to the second by multiplying c by i. ■

Lemma 6.10 *Let $z, c \in \mathbb{C}$, $c \# 0$, $n \in \mathbb{N}$. Then*

$$z^n = c \vee z^n = -c,$$

if the conjunction of the following two equations holds.

$$(|z|^2)^n = |c|^2, \tag{1}$$
$$z^n\bar{c} - \bar{z}^n c = 0. \tag{2}$$

(If $n > 0$, the first determines a circle in the complex plane, while the second determines a number of lines through the origin.)

Proof Given these two equations, $z^n\bar{c} = \bar{z}^n c$, and so

$$(z^n)^2\bar{c} = z^n z^n \bar{c} = z^n \bar{z}^n c = (|z|^2)^n c = |c|^2 c = c^2\bar{c}.$$

Because $c \# 0$ we can divide by $\bar{c}$ and hence $(z^n)^2 = c^2$. Again because $c \# 0$ from this it follows that $z^n = c \vee z^n = -c$. ■

Lemma 6.11 *For $a, b \in \mathbb{R}$, $b \# 0$, $n \in \mathbb{N}$,*

$$\frac{(r+i)^n(a-ib) - (r-i)^n(a+ib)}{2i}$$

is a polynomial in r of degree n with real coefficients.

Proof This is equal to $Im\ (r+i)^n(a-ib)$, so it will be real. Now $(r+i)^n(a-ib)$ clearly has degree n, and because its head coefficient is $a - ib$ and $b \# 0$, its imaginary part will have head coefficient $-b$, and so it also will have degree n. ■

Proposition 6.12 *For $c = a + ib \in \mathbb{C}$, $c \# 0$, and $n \in \mathbb{N}$, n odd, there exists a $z \in \mathbb{C}$ such that $z^n = c$.*

Proof We first treat the case that $b \# 0$. Then by Lemma 6.11, $f(r) \equiv \big((r+i)^n(a-ib) - (r-i)^n(a+ib)\big)/2i$ is a polynomial of odd degree with real coefficients. Hence it has a root in $\mathbb{R}$. We now solve the following two equations in x and y.

$$\begin{aligned} r &= x/y, \\ (x^2+y^2)^n &= a^2+b^2. \end{aligned}$$

From the fact that r is a root of f, by multiplying with $2iy^n$ we find that $(x+iy)^n(a-ib) - (x-iy)^n(a+ib) = 0$, and from Lemma 6.10 then $(x+iy)^n = a+ib \vee (x+iy)^n = -a-ib$. In the first case $z = x+iy$, and in the second case $z = -x-iy$ will be a solution to $z^n = c$.

The case that $a \# 0$ reduces to the other one by multiplying c by i. ■

Theorem 6.13 *For $c \in \mathbb{C}$, $c \# 0$ and $n \in \mathbb{N}^+$ there exists an $z \in \mathbb{C}$ such that $z^n = c$.*

Proof This combines the ability to take square and odd roots. Write n as the product of a power of 2 and an odd factor and iterate taking roots. (Note that this uses strong extensionality of taking powers: we need that the result of taking a root is again $\# 0$.) ■

7 Proof of the Fundamental Theorem of Algebra

Proposition 7.1 (Kneser Lemma) *For every $n \in \mathbb{N}$, $n \geq 2$ there exists a $q \in \mathbb{R}$, $0 < q < 1$ such that for every polynomial over $\mathbb{C}$ with leading coefficient 1*

$$f(x) = x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0$$

one has

$$\forall c > |b_0| \exists z \in \mathbb{C}[|z| < c^{1/n} \wedge |f(z)| < qc].$$

Before proving the Kneser Lemma, we state the so called 'Main Lemma' that gives the main ingredients for proving the Kneser Lemma. The advantage of the Main Lemma is that it is just about real numbers; the complex numbers only come in with the Kneser Lemma. There is a 'Key Lemma' that proves the Main Lemma. We state the Key Lemma first.

Lemma 7.2 (Key Lemma) *For every $n \geq 2$, $\epsilon > 0$ and $a_0, \ldots, a_n \geq 0$ with $a_n = 1$, $a_0 > \epsilon$, there exists*

1. $t > 0$
2. $k_0 \geq k_1, \geq k_2 \geq \ldots,$

such that

$$a_{k_0} t^{k_0} = a_0 + \epsilon$$

and moreover for every j, if we let $k = k_j$ and $r = 3^{-j}t$:

$$a_k r^k > a_i r^i - \epsilon \qquad \text{for all } i \in \{1, \ldots, n\}$$

From the Key Lemma we obtain the Main Lemma

Lemma 7.3 (Main Lemma) *For every $n \geq 2$, $\epsilon > 0$ and $a_0, \ldots, a_n \geq 0$ with $a_n = 1$, $a_0 > \epsilon$, there exists*

1. $k \in \{1, \ldots, n\}$,
2. $r > 0$

such that

$$\begin{aligned} r^n &< a_0, && (3)\\ a_k r^k &< a_0, && (4)\\ 3^{-2n^2} a_0 - 2\epsilon &< a_k r^k, && (5)\\ \sum_{\substack{i=1\\ i\neq k}}^{n} a_i r^i &< (1-3^{-n}) a_k r^k + 3^n \epsilon. && (6) \end{aligned}$$

The Main Lemma is the crucial property about reals to prove the Kneser Lemma.

Proof of the Key Lemma, 7.2 We prove the Key Lemma in a sequence of smaller Lemmata, some specifically related to FTA, some of a more general nature.

Lemma 7.4 *For $n > 0$, $a_1, \ldots, a_n \in \mathbb{R}$ and $\epsilon > 0$ there always is a $k \in \{1, \ldots, n\}$ such that for all $i \in \{1, \ldots, n\}$:*

$$a_k > a_i - \epsilon$$

Proof Induction with respect to n.

Lemma 7.5 *For each sequence $k_0 \geq k_1 \geq k_2 \geq \ldots \in \{1, \ldots, n\}$ there is a $j \in \mathbb{N}$ with $j < 2n$ such that $k_{j-1} = k_j = k_{j+1}$.*

Proof Induction with respect to n.

Lemma 7.6 *Let $n > 0$ and $\epsilon > 0$. Then for every $a_0, \ldots, a_n \geq 0$ with $a_0 > \epsilon$ and $a_n = 1$, there exist $t > 0$ and $k \in \{1, \ldots, n\}$ such that:*

$$a_k t^k = a_0 - \epsilon$$

and such that for all $i \in \{1, \ldots, n\}$:

$$a_i t^i < a_0$$

Proof Start with $k = n$ and $t = \sqrt[n]{a_0 - \epsilon}$. Then consider in turn for i the values $n-1$ down to 1. At each i either $a_i t^i < a_0$ or $a_i t^i > a_0 - \epsilon$ (for the value of t that is current at that time.) In the first case do nothing, but in the second case set k to i and t to $\sqrt[i]{(a_0 - \epsilon)/a_i}$ (in which case t will decrease.) This will give at the end a suitable k and t.

Proof [of the Key Lemma, 7.2] Let $n \geq 2$, $\epsilon > 0$ and $a_0, \ldots, a_n \geq 0$ with $a_n = 1$, $a_0 > 0$ be given. Choose t and k_0 according to Lemma 7.6.

To get k_{j+1} from k_j, let $k = k_j$, $r = 3^j t$ and apply lemma 7.4 with $\epsilon/2$ to the sequence

$$a_1(r/3), a_2(r/3)^2, \ldots, a_k(r/3)^k$$

to get $k' = k_{j+1}$. Then for $i \leq k$ the inequality for k_{j+1} directly follows, while for $i > k$ we have:

$$a_k(r/3)^k = 3^{-k} a_k r^k > 3^{-k}\left(a_i r^i - \epsilon\right) = 3^{-k} a_i r^i - 3^{-k}\epsilon > a_i (r/3)^i - \epsilon/2$$

and so:

$$a_{k'}(r/3)^{k'} > a_k(r/3)^k - \epsilon/2 > a_i(r/3)^i - \epsilon$$

Proof of the Main Lemma, 7.3 We also prove the Main Lemma in a sequence of smaller Lemmata.

Lemma 7.7 *For every $n \geq 2$, $\epsilon > 0$ and $a_0, \ldots, a_n \geq 0$ with $a_n = 1$, $a_0 > \epsilon$, if there exist*

1. $t > 0$
2. $k_0 \geq k_1, \geq k_2 \geq \ldots,$

such that

$$a_{k_0} t^{k_0} = a_0 + \epsilon$$

and moreover for every j, if we let $k = k_j$ and $r = 3^{-j}t$:

$$a_k r^k > a_i r^i - \epsilon \qquad \text{for all } i \in \{1, \ldots, n\}$$

then we have for all j, writing again $k = k_j$ and $r = 3^{-j}t$,

$$\begin{aligned} r^n &< a_0 \\ a_k r^k &< a_0 \\ 3^{-jn} a_0 - 2\epsilon &< a_k r^k \end{aligned}$$

Proof We have $r \leq t$ and so for all i we have $a_i r^i \leq a_i t^i < a_{k_0} t^{k_0} + \epsilon = a_0$. Of this statement $r^n < a_0$ and $a_k r^k < a_0$ are special cases. Finally, from $a_{k_0} r^{k_0} = 3^{-jk_0} a_{k_0} t^{k_0} \geq 3^{-jn} a_{k_0} t^{k_0} = 3^{-jn}(a_0 - \epsilon) > 3^{-jn} a_0 - \epsilon$ it follows that $a_k r^k > a_{k_0} r^{k_0} - \epsilon > 3^{-jn} a_0 - 2\epsilon$.

Lemma 7.8 *For every $n \geq 2$, $\epsilon > 0$ and $a_0, \ldots, a_n \geq 0$ with $a_n = 1$, $a_0 > \epsilon$, if there exist*

1. $t > 0$
2. $k_0 \geq k_1, \geq k_2 \geq \ldots,$

such that for every j, if we let $k = k_j$ and $r = 3^{-j}t$:

$$a_k r^k > a_i r^i - \epsilon \qquad \text{for all } i \in \{1, \ldots, n\}$$

then there is a $j_0 < 2n$ such that, writing $k = k_{j_0}$ and $r = 3^{-j_0}t$,

$$\begin{aligned} a_k (r/3)^k &> a_i (r/3)^i - \epsilon \qquad \text{for all } i \in \{1, \ldots, n\} \\ a_k (3r)^k &> a_i (3r)^i - \epsilon \qquad \text{for all } i \in \{1, \ldots, n\} \end{aligned}$$

Proof From Lemma 7.5 it follows that there is a $j_0 < 2n$ such that $k_{j_0 - 1} = k_{j_0} = k_{j_0+1}$. Writing k for k_{j_0}, it immediately follows from $k_{j_0-1} = k_{j_0}$ and the properties of the k-sequence that $a_k(3r)^k > a_i(3r)^i - \epsilon$. Similarly, it follows from $k_{j_0} = k_{j_0+1}$ and the properties of the k-sequence that $a_k(r/3)^k > a_i(r/3)^i - \epsilon$.

Lemma 7.9 *For every $\epsilon > 0$, $a_1, \ldots, a_n \geq 0$, $k \in \{1, \ldots, n\}$ and $r > 0$ such that for all $i \in \{1, \ldots, n\}$:*

$$a_k (r/3)^k > a_i (r/3)^i - \epsilon$$

holds:

$$\sum_{i=1}^{k-1} a_i r^i < \frac{1}{2}(1 - 3^{-n}) a_k r^k + \frac{1}{2} 3^n \epsilon$$

Proof From the assumption it follows that

$$\begin{aligned} a_i r^i &= 3^i a_i (r/3)^i \\ &< 3^i (a_k (r/3)^k + \epsilon) \\ &< 3^{i-k} a_k r^k + 3^i \epsilon \end{aligned}$$

and therefore

$$\begin{aligned} \sum_{i=1}^{k-1} a_i r^i &< \sum_{i=1}^{k-1} \left(3^{i-k} a_k r^k + 3^i \epsilon\right) \\ &= (\sum_{i=1}^{k-1} 3^{i-k}) a_k r^k + (\sum_{i=1}^{k-1} 3^i)\epsilon \\ &= \frac{1}{2}(1 - 3^{1-k}) a_k r^k + \frac{1}{2}(3^k - 3^1)\epsilon \\ &< \frac{1}{2}(1 - 3^{-n}) a_k r^k + \frac{1}{2} 3^n \epsilon \end{aligned}$$

Lemma 7.10 *For every* $\epsilon > 0$*,* $a_1, \dots, a_n \geq 0$*,* $k \in \{1, \dots, n\}$ *and* $r > 0$ *such that for all* $i \in \{1, \dots, n\}$*:*

$$a_k (3r)^k > a_i (3r)^i - \epsilon$$

holds:

$$\sum_{i=k+1}^{n} a_i r^i < \frac{1}{2}(1 - 3^{-n}) a_k r^k + \frac{1}{2} 3^n \epsilon$$

Proof From the assumption it follows that

$$\begin{aligned} a_i r^i &= 3^{-i} a_i (3r)^i \\ &< 3^{-i} (a_k (3r)^k + \epsilon) \\ &< 3^{k-i} a_k r^k + 3^{-i} \epsilon \end{aligned}$$

and therefore

$$\begin{aligned} \sum_{i=k+1}^{n} a_i r^i &< \sum_{i=k+1}^{n} \left(3^{k-i} a_k r^k + 3^{-i} \epsilon\right) \\ &= (\sum_{i=k+1}^{n} 3^{k-i}) a_k r^k + (\sum_{i=k+1}^{n} 3^{-i})\epsilon \\ &= \frac{3}{2}(3^{-1} - 3^{k-n-1}) a_k r^k + \frac{3}{2}(3^{-k-1} - 3^{-n-1})\epsilon \\ &= \frac{1}{2}(1 - 3^{k-n}) a_k r^k + \frac{1}{2}(3^{-k} - 3^{-n})\epsilon \\ &< \frac{1}{2}(1 - 3^{-n}) a_k r^k + \frac{1}{2} 3^n \epsilon \end{aligned}$$

Lemma 7.11 *For every* $\epsilon > 0$*,* $a_1, \dots, a_n \geq 0$*,* $k \in \{1, \dots, n\}$ *and* $r > 0$ *such that for all* $i \in \{1, \dots, n\}$*:*

$$\begin{aligned} a_k (r/3)^k &> a_i (r/3)^i - \epsilon \\ a_k (3r)^k &> a_i (3r)^i - \epsilon \end{aligned}$$

holds:

$$\sum_{\substack{i=1 \\ i \neq k}}^{n} a_i r^i < (1 - 3^{-n}) a_k r^k + 3^n \epsilon$$

Proof This follows immediately from Lemmata 7.9, 7.10.

Proof [of the Main Lemma, 7.3] Take t and $k_0, k_1, \ldots$ according to the Key Lemma 7.2. According to Lemma 7.8 there is a $j_0 < 2n$ such that for $k = k_{j_0}$ and $r = 3^{-j_0}t$ the premises of Lemma 7.11 hold. Hence inequality (6) of the Main Lemma holds:

$$\sum_{\substack{i=1 \\ i \neq k}}^{n} a_i r^i < (1 - 3^{-n}) a_k r^k + 3^n \epsilon$$

Then inequalities (3), (4) and (5) are given by lemma 7.7 (the inequality $3^{-2n^2} a_0 < 3^{-j_0 n} a_0$ holds because $j_0 < 2n$).

Proof of the Kneser Lemma, Proposition 7.1 We prove the Kneser Lemma in a sequence of steps. Let $n \geq 2$. We will show that

$$q := 1 - \frac{1}{3^{2n^2+n}}$$

is a good choice for q. Let

$$f(x) = x^n + b_{n-1}x^{n-1} + \ldots + b_1 x + b_0$$

be a polynomial over $\mathbb{C}$ and let $c \in \mathbb{R}^+$ be such that $c > |b_0|$. We want to apply the Main Lemma taking $a_i := |b_i|$. However, we don't know if $|b_0| \# 0$. Hence we will approximate b_0 by a $b_0' \# 0$ such that $|b_0 - b_0'|$ is sufficiently small and $|b_0'| < c$. Then we will define the real numbers $a_0, \ldots, a_n$ by $a_0 := |b_0'|$, $a_i := |b_i|$ for $1 \leq i < n$ and $a_n := 1$. Now, for a specific choice of z (with $|z|^n < a_0$) the Main Lemma will give an approximation of $|f(z)|$ in terms of a_0 and hence in terms of c. In particular, it will be shown that $|f(z)| < qc$, with q as above.

Lemma 7.12 *Let $a_0, \ldots, a_n \geq 0$ and $b_0, \ldots, b_n \in \mathbb{C}$ with $a_i = |b_i|$ for $i = 1, \ldots, n$. Furthermore, let $k \in \{1, \ldots, n\}$ and $z \in \mathbb{C}$ with $r = |z|$. Then:*

$$\left|\sum_{i=0}^{n} b_i z^i\right| < \left|b_0 + b_k z^k\right| + \sum_{\substack{i=1 \\ i \neq k}}^{n} a_i r^i$$

Proof Repeated application of the triangle inquality for the complex numbers.

The Main Lemma will take care that the second term on the right hand side of the conclusion of Lemma 7.12 is sufficiently small. To assure that the first term is also small enough, a specific value of z can be chosen in such a way that b_0 and $b_k z^k$ cancel eachother out.

Lemma 7.13 *Given a_0, $a_k > 0$, b_0, b_0', $b_k \in \mathbb{C}$, $k \in \{1, \ldots, n\}$, $r > 0$ and $\eta > 0$ such that:*

$$\begin{aligned} |b_0'| &= a_0 \\ |b_k| &= a_k \\ |b_0 - b_0'| &< \eta \\ a_k r^k &< a_0 \end{aligned}$$

then there exists a $z \in \mathbb{C}$ such that $|z| = r$ and:

$$\left|b_0 + b_k z^k\right| < a_0 - a_k r^k + \eta$$

Proof Take

$$z = r \sqrt[k]{-\frac{a_k}{a_0}\frac{b_0'}{b_k}}$$

Then we have:

$$\Big|-\frac{a_k}{a_0}\frac{b_0'}{b_k}\Big| = \frac{a_k}{a_0}\frac{|b_0'|}{|b_k|} = \frac{a_k a_0}{a_0 a_k} = 1$$

so

$$\left|\sqrt[k]{-\frac{a_k}{a_0}\frac{b_0'}{b_k}}\right| = 1$$

and so $|z| = r$.

Because $a_k r^k < a_0$ we get $\left|a_0 - a_k r^k\right| = a_0 - a_k r^k$ and therefore

$$\begin{aligned}
|b_0' + b_k z^k| &= \Big|b_0' + b_k r^k\big(-\frac{a_k}{a_0}\frac{b_0'}{b_k}\big)\Big| \\
&= \big|\frac{b_0'}{a_0}(a_0 - a_k r^k)\big| \\
&= \frac{|b_0'|}{a_0}|a_0 - a_k r^k| \\
&= a_0 - a_k r^k
\end{aligned}$$

From this it follows that $\left|b_0 + b_k z^k\right| \leq \left|b_0 + b_k z^k\right| + |b_0 - b_0'| < a_0 - a_k r^k + \eta$.

Lemma 7.14 *For $\eta > 0$ and $z \in \mathbb{C}$ there is a $z' \in \mathbb{C}$ with $z' \# 0$ and $|z' - z| < \eta$.*

Proof Because $z + \eta/2 \;\#\; z - \eta/2$, either $z + \eta/2 \;\#\; 0$ or $z - \eta/2 \;\#\; 0$. For both choices $|z' - z| = \eta/2 < \eta$.

Lemma 7.15 *Let be given $b_0 \in \mathbb{C}$ and $c \in \mathbb{R}$ with $|b_0| < c$. Then there are $b_0' \in \mathbb{C}$, a_0 and $\eta > 0$ such that:*

$$\begin{aligned}
|b_0 - b_0'| &< \eta && (7) \\
|b_0'| &= a_0 && (8) \\
a_0 &> 0 && (9) \\
a_0 + 3\eta &< c && (10)
\end{aligned}$$

and an $\epsilon > 0$ such that:

$$\begin{aligned}
2(3^n + 1)\epsilon &< \eta && (11) \\
2\epsilon &< 3^{-2n^2} a_0 && (12) \\
\epsilon &< a_0 && (13)
\end{aligned}$$

Proof Take

$$\eta = \frac{1}{4}(c - |b_0|)$$

so $|b_0| = c - 4\eta$. Then choose an arbitrary $b_0' \# 0$ with $|b_0' - b_0| < \eta$ and take $a_0 = |b_0'|$. To see that (10) is satisfied, calculate:

$$a_0 = |b_0'| \leq |b_0' - b_0| + |b_0| < \eta + c - 4\eta = c - 3\eta$$

The existence of a suitable ϵ then follows easily: take $\epsilon > 0$ smaller then $\min(\frac{\eta}{2(3^n+1)}, \frac{a_0}{2\, 3^{2n^2}})$.

Lemma 7.16 *For:*

$$q = 1 - 3^{-2n^2-n}$$

we have that $q > \frac{1}{2}$ *and because of that inequalities (10) and (11) imply:*

$$qa_0 + 3^n\epsilon + \epsilon + \eta < qc$$

Proof We get

$$a_0 + 2\cdot 3^n\epsilon + 2\epsilon + 2\eta = a_0 + 2(3^n+1)\epsilon + 2\eta < a_0 + \eta + 2\eta < c$$

Using that $1 < 2q$, this gives

$$qa_0 + 3^n\epsilon + \epsilon + \eta < qa_0 + 2q3^n\epsilon + 2q\epsilon + 2q\eta \ = q(a_0 + 2\cdot 3^n\epsilon + 2\epsilon + 2\eta) < qc$$

Proof [of the Kneser Lemma, Proposition 7.1] Take b_0', a_0, η and ϵ as in lemma 7.15. Take $a_i = |b_i|$ for $i \in \{1, \ldots, n\}$. Take r and k as in lemma 7.3. Finally take z as in lemma 7.13.

Then plugging all conditions and results of lemmas 7.3, 7.12, 7.13, 7.15 and 7.16 together we get

$$r^n < a_0 < c - 3\eta < c$$

so

$$|z| = r < c^{1/n}$$

and

$$\begin{aligned}
|\sum_{i=0}^{n} b_i z^i| &< |b_0 + b_k z^k| + \sum_{\substack{i=1\\ i\neq k}}^{n} a_i r^i \\
&< \left(a_0 - a_k r^k + \eta\right) + \left((1-3^{-n})a_k r^k + 3^n\epsilon\right) \\
&= a_0 - 3^{-n} a_k r^k + 3^n\epsilon + \eta \\
&< a_0 - 3^{-n}(3^{-2n^2} a_0 - 2\epsilon) + 3^n\epsilon + \eta \\
&= (1 - 3^{-2n^2-n})a_0 + 3^n\epsilon + 3^{-n}2\epsilon + \eta \\
&< (1 - 3^{-2n^2-n})a_0 + 3^n\epsilon + \epsilon + \eta \\
&= qa_0 + 3^n\epsilon + \epsilon + \eta \\
&< qc
\end{aligned}$$

Fundamental Theorem for regular polynomials

Proposition 7.17 *Let* $f(x) = x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$*, with* $a_i \in \mathbb{C}$*. Then for some* $z \in \mathbb{C}$ *one has* $f(z) = 0$*.*

Proof Let $c \in \mathbb{R}^+$ with $c > |a_0|$. We will construct a Cauchy sequence $z_i \in \mathbb{C}$ such that for all m

1. $|f(z_m)| < q^m c$
2. $|z_{m+1} - z_m| \leq (q^m c)^{1/n}$

for some $q \in (0,1)$. Then $z = \lim_{i\to\infty} z_i$ exists and by continuity of f one has

$$|f(z)| = \lim_{i\to\infty} |f(z_i)| \leq \lim_{i\to\infty} q^i c = 0,$$

so $f(z) = 0$.

Now, if 1 and 2 are satisfied, then indeed the z_i form a Cauchy seqence:

$$\begin{aligned}
|z_{m+k} - z_m| &\leq |z_{m+k} - z_{m+k-1}| + \ldots + |z_{m+1} - z_m| \\
&\leq (q^{\frac{m+k-1}{n}} + q^{\frac{m+k-2}{n}} + \ldots + q^{\frac{m}{n}})c^{1/n} \\
&= \frac{q^{\frac{m}{n}} - q^{\frac{m+k}{n}}}{1 - q^{1/n}} c^{1/n} \\
&= q^{\frac{m}{n}} \frac{1 - q^{\frac{k}{n}}}{1 - q^{1/n}} c^{1/n} \\
&\leq q^{\frac{m}{n}} \frac{c^{1/n}}{1 - q^{1/n}}.
\end{aligned}$$

By choosing m sufficiently large (n is fixed), this last expression can be made arbitrarily small.

The construction of z_i is as follows. Take $z_0 = 0$. Then indeed $|f(z_0)| = |f(0)| < q^0 c$. Now suppose z_m is defined satisfying 1. Apply the Kneser Lemma to f_{z_m} where

$$f_{z_m}(x) = f(x + z_m)$$

and taking $q^m c$ for c. (Note that f_{z_m} has the same degree as f.) We obtain a z such that

$$|z| < (q^m c)^{1/n} \wedge |f_{z_m}(z)| < q^{m+1} c.$$

Now take $z_{m+1} = z + z_m$. Then we have that 1 is valid: $|f(z_{m+1})| = |f(z + z_m)| = |f_{z_m}(z)| < q^{m+1} c$.

Moreover, we also have 2: $|z_{m+1} - z_m| = |z| < (q^m c)^{1/n}$. ■

Corollary 7.18 *1. Every regular polynomial $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$ over $\mathbb{C}$ has a root.*

2. Moreover, such f can be factorized as follows

$$\overline{f}(x) = a_n(x - \alpha_1)\ldots(x - \alpha_n).$$

Proof 1. Divide f by a_n to obtain a polynomial g with leading coefficient 1, satisfying $a_n g(x) = f(x)$. Then any root of g is a root of f.

2. If α_1 is a root of f, then

$$f(x) = (x - \alpha_1) f_{n-1}(x),$$

by the Remainder Theorem (4.11) with f_{n-1} being equal to $a_n x^{n-1} + \ldots$, hence also regular. By (i) f_{n-1} has a root α_2 and hence

$$f_{n-1}(x) = (x - \alpha_2) f_{n-2}(x).$$

Continuing this way one obtains

$$f(x) = (x - \alpha_1)\ldots(x - \alpha_n)a_n. \blacksquare$$

Proposition 7.19 *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, with $a_i \in \mathbb{C}$. Suppose that $a_k \# 0$, for some $0 < k \leq n$. Then f can be factorized as follows in linear factors.*

$$\overline{f}(x) = (\beta_1 x - \delta_1)\ldots(\beta_n x - \delta_n).$$

Proof Let f be given. Take a $c \in \mathbb{C}$ as in Lemma 4.14 such that $f(c) \# 0$ and let $g = (f_c)^{\tilde{}}$. Note that by Lemma 4.17 we have

$$\overline{g}(y) = y^n \overline{f}(c + \frac{1}{y}), \text{ for } y \# 0.$$

satisfying

$$\overline{g}(y) = \overline{f}(c)y^n + \frac{\overline{f'}(c)}{1!}y^{n-1} + \ldots + \frac{\overline{f^{(n)}}(c)}{n!}.$$

It follows that g is regular, with leading coefficient $\overline{f}(c)$, and hence by the corollary g is a product

$$\overline{g}(y) = \overline{f}(c)(y-\alpha_1)\ldots(y-\alpha_n). \tag{1}$$

Now, for $x \# c$,

$$\begin{aligned}
\overline{f}(x) &= \overline{g}(\frac{1}{x-c})(x-c)^n \\
&= \overline{f}(c)(\frac{1}{x-c}-\alpha_1)\ldots(\frac{1}{x-c}-\alpha_n)(x-c)^n \\
&= \overline{f}(c)(1-\alpha_1(x-c))\ldots(1-\alpha_n(x-c)) \\
&= \overline{f}(c)(1+c-\alpha_1 x)\ldots(1+c-\alpha_n x).
\end{aligned}$$

So, we are done by taking $\beta_i := -\overline{f}(c)\alpha_i$ and $\delta_i := -\overline{f}(c)(1+c)$. ■

Theorem 7.20 *Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$, with $a_i \in \mathbb{C}$. Suppose that $a_k \# 0$, for some $0 < k \leq n$. Then f can be factorized as follows in linear factors.*

$$\overline{f}(x) = (x-\alpha_1)\ldots(x-\alpha_k)(\beta_{k+1}x-\delta_{k+1})\ldots(\beta_n x-\delta_n).$$

So, in particular, f has k zeros.

Proof From Proposition 7.19 we conclude that

$$\overline{f}(x) = (\beta_1 x-\delta_1)\ldots(\beta_n x-\delta_n),$$

for some β_i and δ_i in $\mathbb{C}$. We prove by induction on n that at least k of the β_i can be chosen apart from 0; hence, by dividing out these factors, we obtain the α_i, β_i and δ_i of the statement.

$n = 1$ Then $f(x) = a_1 x + a_0$ and $a_1 \# 0$, so we are done.

$n+1$ Now $\overline{f}(x) = (\beta_1 x-\delta_1)\ldots(\beta_{n+1}x-\delta_{n+1})$. Writing $h(x) = (\beta_2 x-\delta_2)\ldots(\beta_{n+1}x-\delta_{n+1})$, we find that h is a polynomial of length n, say $h(x) = h_n x^n + \ldots + h_0$. We find that

$$a_k = \beta_1 h_{k-1} - \delta_1 h_k.$$

As $a_k \# 0$, we conclude that $\beta_1 h_{k-1} \# 0$ or $\delta_1 h_k \# 0$. In the first case, $\beta_1 \# 0$ and from the induction hypothesis we derive that $k-1$ from the $\beta_2,\ldots,\beta_n$ are $\# 0$, so we are done. In the second case, we conclude from the induction hypothesis that k from the $\beta_2,\ldots,\beta_n$ are $\# 0$, so we are done.

■

References

[1] Heyting, *Intuitionism, an introduction*, North-Holland, 1956, 133 pp.

[2] J.E. Littlewood, Every polynomial has a root, Journal of the London Math. Soc. 16 (1941), pp. 95 – 98.

[3] Ruitenburg, *Intuitionistic Algebra, Theory and Sheaf Models*, Ph.D. Thesis, Utrecht University, June 1982, 143 pp.

[4] Mines, Richman and Ruitenburg, *A Course in Constructive Algebra*, Universitext, Springer-Verlag, 344 pp.

[5] Troelstra and van Dalen, *Constructivism in Mathematics, an Introduction, Vol 1*, volume 121 in Studies in Logica and The Foundations of Mathematics, North-Holland, 1988, 342 pp.

[6] Troelstra and van Dalen, *Constructivism in Mathematics, an Introduction, Vol 2*, volume 123 in Studies in Logica and The Foundations of Mathematics, North-Holland, 1988, 879 pp.