Appendix A: Notes

p. 5: We read "$s \in A$" as "s is a member of A", "s belongs to A", or "A contains s". "$A \subset B$" (equivalently, "$B \supset A$") is read "A is a subset of B", "A is included in B", or "B includes A". We say that "A is strictly included in B" or "B strictly includes A" if A is a subset of B and $A \neq B$; *i.e.*, $A \subsetneqq B$.

p. 5: The symbol "$\Longrightarrow$" denotes logical implication. Thus, we interpret "$P \implies Q$" as "if P, then Q", or "P only if Q".

p. 6: "$:\Longleftrightarrow$" is to be interpreted "means, by definition, that". Thus, defining the relation ρ on $\mathbb{N}$ by

$$m \,\rho\, n \;:\Longleftrightarrow\; n = m + 1$$

for all $m, n \in \mathbb{N}$ says that $m \,\rho\, n$ means that $n = m + 1$ whenever $m, n \in \mathbb{N}$.

p. 7: "$:=$" is to be interpreted as "is equal, by definition, to".

p. 7: We read "$\{z \in \mathcal{E} \mid z \,\rho\, x\}$" as "the set of all members z of $\mathcal{E}$ such that $z \,\rho\, x$ is valid". This notation is useful for describing subsets of a given set whose members satisfy specific conditions.

p. 8: We denote by "$\cap$" and "$\cup$" the usual set intersection and set union. If I is some index set, and if $(S_i \mid i \in I)$ is a family of sets indexed on I, then we denote by "$\bigcap_{i \in I} S_i$" and "$\bigcup_{i \in I} S_i$" the intersection and union, respectively, of the sets in the given family. If $\mathcal{C}$ is a collection of sets, we denote by "$\bigcap \mathcal{C}$" and "$\bigcup \mathcal{C}$" the intersection and union, respectively, of all members of $\mathcal{C}$. Thus, considering $\mathcal{C}$ as a family $(S \mid S \in \mathcal{C})$ indexed on $\mathcal{C}$, we have

$$\bigcap \mathcal{C} = \bigcap_{S \in \mathcal{C}} S \quad \text{and} \quad \bigcup \mathcal{C} = \bigcup_{S \in \mathcal{C}} S.$$

p. 8: We use "$\emptyset$" as a symbol for the empty set.

p. 10: We denote logical equivalence by the symbol "$\Longleftrightarrow$". Thus, we interpret "$P \Longleftrightarrow Q$" as "P if and only if Q".

p. 14: We use the symbol "$\mathbb{R}$" to denote the set of all real numbers, as well as the symbols "$<$", "$\leq$", "$>$", and "$\geq$" to denote the usual relations on $\mathbb{R}$. We denote by "$\mathbb{P}$" the set of all positive reals; that is, $\mathbb{P} := \{r \in \mathbb{R} \,|\, r \geq 0\}$. We denote by "$\mathbb{P}^{\times}$" the set of all strictly positive reals; that is, we have that $\mathbb{P}^{\times} := \{r \in \mathbb{R} \,|\, r > 0\}$. We denote by "$\mathbb{N}$" the *natural numbers;* that is, the subset of $\mathbb{P}$ consisting of the integers which are members of $\mathbb{P}$. We denote by $\mathbb{N}^{\times}$ the set of all natural numbers excluding 0.

p. 14, 15: For completeness, we give the following notations for various types of intervals in $\mathbb{R}$.

$$\begin{aligned} [a,b[\; &:= [a,b] \setminus \{b\}, &]a,\infty[\; &:= [a,\infty[\setminus \{a\}, \\ ]a,b] &:= [a,b] \setminus \{a\}, &]-\infty,b] &:= \{c \in \mathbb{R} \,|\, c \leq b\}, \\ ]a,b[\; &:= [a,b] \setminus \{a,b\}, &]-\infty,b[\; &:= \;]-\infty,b] \setminus \{b\}, \\ [a,\infty[\; &:= \{c \in \mathbb{R} \,|\, a \leq c\}, &]-\infty,\infty[\; &:= \mathbb{R}. \end{aligned}$$

(See Note for p. 20 about "$\setminus$".)

p. 19: If $\mathcal{E}$ is a set, and $\mathsf{P} \subset \mathsf{Sub}\,\mathcal{E}$ (see Note for p. 30) is such that $\emptyset \notin \mathsf{P}$, $\bigcup \mathsf{P} = \mathcal{E}$, and for all $\mathcal{S}, \mathcal{T} \in \mathsf{P}$, $\mathcal{S} \neq \mathcal{T} \implies \mathcal{S} \cap \mathcal{T} = \emptyset$, then P is said to be a *partition of* $\mathcal{E}$.

p. 20: The symbol "$\setminus$" denotes set-difference. Thus,

$$A \setminus B := \{x \in A \,|\, x \notin B\}.$$

p. 21: See the Note for p. 5 for a description of "$\subsetneqq$".

p. 21: By $f : D \to C$, we mean a mapping (sometimes referred to as a *function*) which assigns to each member of D a member of C. D is called the *domain of* f and C is called the *codomain of* f.

p. 22: The range of $w_{\mathcal{L}}$ is denoted by $\mathsf{Rng}\, w_{\mathcal{L}} := \{w_{\mathcal{L}}(\sigma) \,|\, \sigma \in \Lambda_{\mathcal{L}}\}$.

p. 25: We define $\mathbb{R}^2$ by

$$\mathbb{R}^2 := \{(x,y) \,|\, x, y \in \mathbb{R}\}.$$

In an analogous way, when $n \in \mathbb{N}$ and $n \geq 3$, we define $\mathbb{R}^n$ to be the set of all lists of length n whose terms are in $\mathbb{R}$.

p. 30: By "$I \times S$", we mean the set of all pairs whose first term belongs to I and whose second term belongs to S. Thus, $I \times S = \{(i, s) \mid i \in I, s \in S\}$.

p. 30: "$\mathsf{Sub}\,\mathcal{E}$" denotes the set of all subsets of $\mathcal{E}$.

p. 39: See the Note for p. 14 for a description of "$\mathbb{P}$".

p. 41: If A is a nonempty subset of $\mathbb{R}$ bounded above (that is, there is $M \in \mathbb{R}$ such that $a \leq M$ for all $a \in A$), we denote by "$\sup A$" the *supremum of A;* that is, the least upper bound of A. Thus, if $c \in \mathbb{R}$ satisfies $a \leq c$ for all $a \in A$, then necessarily $\sup A \leq c$. Likewise, if A is bounded below, we denote by "$\inf A$" the *infimum of A;* that is, the greatest lower bound of A. See an introductory text on real analysis for more details.

p. 42: "$\#\lambda$" denotes the cardinality of the set λ; *e.g.*, $\#\{\{1\}, 4, \{3, 5\}\} = 3$.

p. 42: "$m..n$" denotes the set $\{c \in \mathbb{N} \mid m \leq c \leq n\}$. Thus, $1..4 = \{1, 2, 3, 4\}$. We have, from the definition, that $m > n$ implies that $m..n = \emptyset$.

p. 42: If $\bar{\lambda}$ is a list of length m, we denote by "$\mathsf{Rng}\,\bar{\lambda}$" the *range of* $\bar{\lambda}$*;* that is, $\mathsf{Rng}\,\bar{\lambda} = \{\bar{\lambda}_i \mid i \in 1..m\}$. $\mathsf{Rng}\,\bar{\lambda}$ is simply the *set* of all terms in the *list* $\bar{\lambda}$.

p. 42: Often the sum "$\sum_{i \in m..n} a_i$" is written "$\sum_{i=m}^{n} a_i$".

p. 43: See the Note for p. 41 for a description of "inf".

p. 44: See the Note for p. 14 for a description of $\mathbb{N}$.

p. 51: Let a mapping $f : D \to C$ be given. f is said to be *injective* if $f(x) = f(y) \implies x = y$ for all $x, y \in D$. f is said to be *surjective* if for all $z \in C$, there is some $x \in D$ such that $f(x) = z$. f is said to be *bijective* (or *invertible*) if f is both injective and surjective.

p. 52: We mean by "$I + a$" the set $\{i + a \mid i \in I\}$.

p. 57: When f is invertible, we denote by "$f^{\leftarrow}$" the inverse of f. Thus, if $f : D \to C$, then for all $x \in D$ we have $f^{\leftarrow}(f(x)) = x$, and for all $y \in C$ we have $f(f^{\leftarrow}(y)) = y$.

p. 58: If $f : D \to C$ and $E \subset D$, we write "$f_{>}(E)$" for the *image of E under f*, given by $f_{>}(E) := \{f(e) \mid e \in E\}$.

p. 58: We denote by "$\circ$" the usual composition of mappings.

p. 60: By "$(t \mapsto (f(t), g(t))) : I \to \mathbb{R}^2$" we mean the mapping $p : I \to \mathbb{R}^2$ given by $p(t) := (f(t), g(t))$ for all $t \in I$. In this case, the notation is used to avoid introducing a new symbol for a mapping which is referred to only once.

p. 66: By $1_{\mathcal{S}}$, we mean the identity mapping on $\mathcal{S}$. Thus, $1_{\mathcal{S}} : \mathcal{S} \to \mathcal{S}$ satisfies $1_{\mathcal{S}}(x) = x$ for all $x \in \mathcal{S}$.

p. 72: By "$\mathcal{V}^{\times}$", we mean $\mathcal{V} \setminus \{\mathbf{0}\}$. The symbol "$\times$" as a superscript denotes the removal of the zero (relative to the set under consideration). See the note for p. 14 about "$\mathbb{R}$" for an analogous use of "$\times$" as a superscript.

p. 86: Given a mapping $f : D \to C$ and $S \subset D$, we define

$$f_{>}(S) := \{f(x) \mid x \in S\}.$$

Analogously, given $T \subset C$, we define

$$f^{<}(T) := \{x \in D \mid f(x) \in T\}.$$

p. 88: See the note for p. 8 concerning "$\bigcap$".

p. 98: Given a mapping $f : D \to C$ and a subset E of D, we denote by $f|_E$ the mapping $f|_E : E \to C$ given by $f|_E(x) := f(x)$ for all $x \in E$. The mapping $f|_E$ is obtained from the mapping f by restricting the domain to E.

p. 102: See the note for p. 8 concerning "$\bigcup \mathsf{F}$".

p. 145: We use "$\approx$" to mean "approximately equal to". Thus, we have $\pi \approx 3.1416$.

p. 179: We use "$\ll$" to mean "very much less than". Thus, we would have both $0.00001 \ll 1$ and $1 \ll 1{,}000{,}000$.

Appendix B: Abbreviations and Symbols

Appendix C: Relations

Consider the statement “Jack is a child of Jill”. We may contrive similar sentences, such as “John is a child of Susan” or “P_1 is a child of P_2”, where P_1 and P_2 represent arbitrary people. Although not all such sentences are necessarily true (since Susan might be John’s older sister), each sentence is meaningful; that is, it makes sense to ask whether each sentence is true. We say that “is a child of” describes a *relation* on the set of all people. In other words, given any two people P_1 and P_2, it makes sense to ask, “Is P_1 a child of P_2?”

C01 Definition: *We say that* ρ *describes a* **relation on a given set** $\mathcal{D}$ *if* $x \,\rho\, y$ *is a meaningful statement for all* $x, y \in \mathcal{D}$. *In this case,* $\mathcal{D}$ *is called the* **domain** *of the relation.*

A frequently encountered relation on a set $\mathcal{D}$ is the **equality relation**, $=_{\mathcal{D}}$, described by

$$x =_{\mathcal{D}} y :\Longleftrightarrow x = y$$

for all $x, y \in \mathcal{D}$.

Let a set $\mathcal{D}$ and a relation ρ on $\mathcal{D}$ be given.

C02 Definition: *We define the* **graph of** ρ *by*

$$\mathsf{Gr}(\rho) := \{(x, y) \in \mathcal{D} \times \mathcal{D} \,|\, x \,\rho\, y\}.$$

For example, if we are given the relation “less than or equal to” on $\mathbb{R}$ (symbolized by $\leq_{\mathbb{R}}$), then the graph of $\leq_{\mathbb{R}}$ is the set of pairs (x, y) with $x \leq_{\mathbb{R}} y$. We remark that we may also speak of the “less than or equal to” relation on other sets, such as $\mathbb{N}$; such a relation would be symbolized by $\leq_{\mathbb{N}}$. Because it is simpler to write and has become conventional, the symbol “$\leq$” is often used to represent both relations. When the “$\leq$” symbol is used in this way, it is important to specify the domain of the relation so that no ambiguity arises, such as, “Consider the relation $\leq$ on $\mathbb{R}$...”. The same convention is used for the “less than” relation (symbolized by $<$) on a given set.

C03 Definition: *For any subset* $\mathcal{S}$ *of* $\mathcal{D}$, *we define the* **restriction of** ρ **to** $\mathcal{S}$ *to be the relation* $\rho|_{\mathcal{S}}$ *on* $\mathcal{S}$ *such that for all* $x, y \in \mathcal{S}$, *we have*

$$x \,\rho|_{\mathcal{S}}\, y :\Longleftrightarrow x \,\rho\, y.$$

We note that $\mathsf{Gr}(\rho|_{\mathcal{S}}) = \mathsf{Gr}(\rho) \cap (\mathcal{S} \times \mathcal{S})$.

For example, consider the relation $\leq$ on $\mathbb{R}$. Then $\leq|_{\mathbb{N}}$ is a relation on $\mathbb{N}$, and $\mathsf{Gr}(\leq|_{\mathbb{N}})$ is the set of pairs of natural numbers (p, q) such that $p \leq q$; this is the same as $\mathsf{Gr}(\leq) \cap (\mathbb{N} \times \mathbb{N})$. Note that $\frac{3}{4} \leq|_{\mathbb{N}} \pi$ does not make sense since $\frac{3}{4}, \pi \notin \mathbb{N}$. Note that the symbols $\leq|_{\mathbb{N}}$ and $\leq_{\mathbb{N}}$ represent the same relation; that is, for all $m, n \in \mathbb{N}$, $m \leq|_{\mathbb{N}} n$ if and only if $m \leq_{\mathbb{N}} n$. The difference is that the former is derived as a restriction of another relation.

C04 Definition: *We say that* ρ *is:*

1. **reflexive** *if for each* $x \in \mathcal{D}$, *we have* $x \,\rho\, x$,
2. **irreflexive** *if for each* $x \in \mathcal{D}$, $x \,\rho\, x$ *is false,*
3. **symmetric** *if for all* $x, y \in \mathcal{D}$, *we have*
$$x \,\rho\, y \implies y \,\rho\, x,$$
4. **antisymmetric** *if for all* $x, y \in \mathcal{D}$, *we have*
$$x \,\rho\, y \text{ and } y \,\rho\, x \implies x = y,$$
5. **strictly antisymmetric** *if for all* $x, y \in \mathcal{D}$, *we have*
$$x \,\rho\, y \implies \textit{not } (y \,\rho\, x),$$
6. **transitive** *if for all* $x, y, z \in \mathcal{D}$, *we have*
$$x \,\rho\, y \text{ and } y \,\rho\, z \implies x \,\rho\, z,$$
7. **total** *if for all* $x, y \in \mathcal{D}$, *we have* $x = y$ *or* $x \,\rho\, y$ *or* $y \,\rho\, x$.

We note that ρ is strictly antisymmetric if and only if it is both antisymmetric and irreflexive. Also, an irreflexive relation ρ is total if and only if for all $x, y \in \mathcal{D}$, we have $x \,\rho\, y$ or $y \,\rho\, x$. Finally, we note that any transitive and irreflexive relation is strictly antisymmetric.

Examples.

1. The relation $\leq$ on $\mathbb{N}$ is a reflexive relation; every natural number is less that or equal to itself.

2. The relation $<$ on $\mathbb{N}$ and the relation "is the father of" on the set of people are irreflexive; no number can be less than itself and no one can be his/her own father.

3. The relation "is a sibling of" on the set of people is a symmetric relation; if John is a sibling of Susan, then certainly Susan is a sibling of John.

4. The relation $\leq$ on $\mathbb{R}$ is antisymmetric; if both $r \leq s$ and $s \leq r$, then it must be the case that $r = s$. The relation $<$ on $\mathbb{R}$ is strictly antisymmetric; if $r < s$, then $s < r$ must be false.

5. The relation "is older than" on the set of all people is a transitive relation.

6. The relation $\leq$ on $\mathbb{R}$ and the relation "is not older than" on the set of all people are total. The "divides" ("is a factor of") relation on $\mathbb{N}^{\times}$, symbolized by div, is not total; neither $7 \,\mathsf{div}\, 3$ nor $3 \,\mathsf{div}\, 7$ are true statements.

C05 Definition: *We say that ρ is an* **order** *if ρ is reflexive, antisymmetric, and transitive. If ρ is strictly antisymmetric and transitive, we say that ρ is a* **strict-order**. *If ρ is a [strict-]order and is also total, we say that ρ is a* **total [strict-]order**.

If $\mathcal{S}$ is a subset of $\mathcal{D}$, we say that $\mathcal{S}$ is **totally ordered with respect to** ρ *if $\rho|_{\mathcal{S}}$ is a total order. If the context is clear, we often simply say that $\mathcal{S}$ is* **totally ordered**. *Note that in this case, ρ itself need not be an order (see an example below).*

For example, div on $\mathbb{N}^\times$ is an order which is not total. However, the subset $\{2^n \,|\, n \in \mathbb{N}\}$ is totally ordered with respect to div. The relation $\leq$ on $\mathbb{R}$ is a total order. Since $\leq$ on $\mathbb{R}$ is total, then any subset of $\mathbb{R}$ is totally ordered with respect to $\leq$.

If ρ is the relation on $\mathbb{R}^2$ defined by

$$(\alpha_1, \beta_1)\,\rho\,(\alpha_2, \beta_2) \;:\Longleftrightarrow\; \alpha_1 \leq \alpha_2$$

for all $(\alpha_1, \beta_1), (\alpha_2, \beta_2) \in \mathbb{R}^2$, we see that ρ is total, although not an order. However, we see that the set $\{(\alpha, \alpha) \,|\, \alpha \in \mathbb{R}\}$ is totally ordered with respect to ρ.

C06 Definition: *For all $x, y \in \mathcal{D}$, we define*

$$[\![x, y]\!]_\rho := \{z \in \mathcal{D} \,|\, x\,\rho\,z \text{ and } z\,\rho\,y\}.$$

This set is called the **ρ-interval between x and y**. *When there is no ambiguity, we often write $[\![x, y]\!]$ for $[\![x, y]\!]_\rho$ and refer to this set as the* **interval between x and y**.

For example, consider the relation $\leq$ on $\mathbb{R}$. Then for $x, y \in \mathbb{R}$, $[\![x, y]\!]_\leq$ is the closed interval $[x, y]$ between x and y. Considering the relation div on $\mathbb{N}^\times$, then $[\![1, n]\!]_{\mathsf{div}}$ is the set of all factors of n. Considering the relation $\leq$ on $\mathbb{N}$, we see that $[\![1, n]\!]_\leq$ is the set of numbers $\{k \in \mathbb{N} \,|\, 1 \leq k \leq n\}$ (the numbers from 1 to n inclusive). The following notation is used to describe such a set:

$$1..n := \{k \in \mathbb{N} \,|\, 1 \leq k \leq n\}.$$

C07 Definition: *Let a subset $\mathcal{S}$ of $\mathcal{D}$ be given. We say that $x \in \mathcal{S}$ is* **minimal in $\mathcal{S}$** *if*

$$\text{For all } y \in \mathcal{S},\;\; y\,\rho\,x \implies y = x.$$

We say that $x \in \mathcal{S}$ is **maximal in $\mathcal{S}$** *if*

$$\text{For all } y \in \mathcal{S},\;\; x\,\rho\,y \implies x = y.$$

We say that $x \in \mathcal{S}$ is a **minimum [maximum] of $\mathcal{S}$** *if for each $y \in \mathcal{S}$, we have $x\,\rho\,y$ [$y\,\rho\,x$].*

If the set $\{x \in \mathcal{D} \,|\, x\,\rho\,s \text{ for all } s \in \mathcal{S}\}$ [$\{x \in \mathcal{D} \,|\, s\,\rho\,x \text{ for all } s \in \mathcal{S}\}$] has a maximum [minimum], we say that this maximum [minimum] is an **infimum [supremum] of $\mathcal{S}$**.

For example, suppose we are given the relation div on $\mathbb{N}^{\times}$. For each $n \in \mathbb{N}^{\times}$, n is maximal in $1..n$. Each prime is minimal in the set $2+\mathbb{N} = \{n \in \mathbb{N} \mid n \geq 2\}$. Thus a set may have more than one minimal (or maximal) element.

Now suppose that we are given a relation ρ on $\mathcal{D}$. Then if $x, y \in \mathcal{D}$ are such that $x \,\rho\, y$, then x is a minimum of $[\![x,y]\!]_\rho$ and y is a maximum of $[\![x,y]\!]_\rho$. It should be noted that when ρ is an order, every subset of $\mathcal{D}$ can have at most one minimum or at most one maximum. In this case, if there is a minimum [maximum], then that minimum [maximum] is the only minimal [maximal] element in the set. Also, in this case, we use the notation $\min \mathcal{S}$ [$\max \mathcal{S}$] for *the* minimum [maximum]. Similarly, we use the notation $\inf \mathcal{S}$ [$\sup \mathcal{S}$] for the infimum [supremum] of $\mathcal{S}$ if there is exactly one such.

Note: The concepts in **Defs. C06** and **C07** are used primarily when ρ is an order. The concept of an interval is occasionally used when ρ is only transitive and not an order.

C08 Proposition: *Let $\prec$ be a transitive relation with domain $\mathcal{D}$. Suppose that $\prec$ is reflexive and total. Then for all $x, y, z \in \mathcal{D}$ such that $x \prec y$ and $y \prec z$, we have*

$$[\![x,y]\!] \cup [\![y,z]\!] = [\![x,z]\!].$$

C09 Definition: *Let ρ and δ be two relations on a set $\mathcal{D}$. We say that ρ* **is finer than** *δ (or equivalently, δ* **is coarser than** *ρ), if*

$$x \,\rho\, y \implies x \,\delta\, y$$

for all $x, y \in \mathcal{D}$. If ρ is finer than δ but is not the same relation as δ, then we say that ρ is **strictly** *finer than δ (equivalently, δ is* **strictly** *coarser that ρ).*

It is clear that ρ is finer than δ if and only if $\mathsf{Gr}(\rho) \subset \mathsf{Gr}(\delta)$. As an example, consider the relations $<$ and $\leq$ on $\mathbb{R}$. Then $<$ is strictly finer than $\leq$, while $\leq$ is strictly coarser than $<$.

C10 Definition: *If δ is a relation on $\mathcal{D}$, we define the relation $\overset{\delta}{=}$ on $\mathcal{D}$ by*

$$x \overset{\delta}{=} y \;:\Longleftrightarrow\; x \,\delta\, y \text{ or } x = y$$

for all $x, y \in \mathcal{D}$. *We also define the relation* $\overset{\delta}{\neq}$ *on* $\mathcal{D}$ *by*

$$x \overset{\delta}{\neq} y \;:\Longleftrightarrow\; x\,\delta\,y \text{ and } x \neq y$$

for all $x, y \in \mathcal{D}$.

For example, we often use "$\leq$" as an abbreviation for "$\overset{\leq}{=}$" on $\mathbb{R}$. Note that if δ on $\mathcal{D}$ is antisymmetric, then $\overset{\delta}{\neq}$ is strictly antisymmetric; while if δ is strictly antisymmetric, then $\overset{\delta}{=}$ is antisymmetric. Thus, if δ is an order on $\mathcal{D}$, then $\overset{\delta}{\neq}$ is a strict-order; while if δ is a strict-order on $\mathcal{D}$, then $\overset{\delta}{=}$ is an order.

C11 Definition: *For a relation* ρ *on* $\mathcal{D}$, *we define the* **reverse of** ρ, *denoted by* $\widetilde{\rho}$, *by*

$$x\,\widetilde{\rho}\,y \;:\Longleftrightarrow\; y\,\rho\,x$$

for all $x, y \in \mathcal{D}$.

We see that the greater than or equal to relation on $\mathbb{R}$, $\geq$, is the reverse of $\leq$ on $\mathbb{R}$.

C12 Definition: *Let* $\sim$ *be a relation on* $\mathcal{D}$. *We say that* $\sim$ *is an* **equivalence relation** *if* $\sim$ *is reflexive, symmetric, and transitive.*

Now let $\sim$ be an equivalence relation on $\mathcal{D}$. For each $x \in \mathcal{D}$, x determines an equivalence class

$$[\![x]\!] := [\![x, x]\!]_{\sim}.$$

The set of all equivalence classes in $\mathcal{D}$,

$$\{[\![x]\!] \mid x \in \mathcal{D}\},$$

is a partition of $\mathcal{D}$ (see the Note for p. 19 in Appendix A concerning partitions).

Appendix D: Linear Spaces

In order to provide a summary of those aspects of linear algebra which are germane to the study of special relativity and to familiarize the reader with our notation and terminology, the following basic Definitions and Propositions are provided.

D01 Definition: *A* **linear space** *is a set $\mathcal{V}$ endowed with structure by the prescription of*

1. *an operation* $\mathsf{add} : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$, *called the* **addition** *in* $\mathcal{V}$,
2. *an operation* $\mathsf{sm} : \mathbb{R} \times \mathcal{V} \to \mathcal{V}$, *called the* **scalar multiplication** *in* $\mathcal{V}$,
3. *an element* $\mathbf{0} \in \mathcal{V}$, *called the* **zero** *of* $\mathcal{V}$, *and*
4. *a mapping* $\mathsf{opp} : \mathcal{V} \to \mathcal{V}$, *called the* **opposition** *in* $\mathcal{V}$,

provided that the following axioms are satisfied for all $\xi, \eta \in \mathbb{R}$ *and* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$:

$$
\begin{array}{lrcl}
(A_1) & \mathsf{add}(\mathbf{u}, \mathsf{add}(\mathbf{v}, \mathbf{w})) & = & \mathsf{add}(\mathsf{add}(\mathbf{u}, \mathbf{v}), \mathbf{w}), \\
(A_2) & \mathsf{add}(\mathbf{u}, \mathbf{v}) & = & \mathsf{add}(\mathbf{v}, \mathbf{u}), \\
(A_3) & \mathsf{add}(\mathbf{u}, \mathbf{0}) & = & \mathbf{u}, \\
(A_4) & \mathsf{add}(\mathbf{u}, \mathsf{opp}(\mathbf{u})) & = & \mathbf{0}, \\
 & & & \\
(S_1) & \mathsf{sm}(\xi, \mathsf{sm}(\eta, \mathbf{u})) & = & \mathsf{sm}(\xi\eta, \mathbf{u}), \\
(S_2) & \mathsf{sm}(\xi + \eta, \mathbf{u}) & = & \mathsf{add}(\mathsf{sm}(\xi, \mathbf{u}), \mathsf{sm}(\eta, \mathbf{u})), \\
(S_3) & \mathsf{sm}(\xi, \mathsf{add}(\mathbf{u}, \mathbf{v})) & = & \mathsf{add}(\mathsf{sm}(\xi, \mathbf{u}), \mathsf{sm}(\xi, \mathbf{v})), \\
(S_4) & \mathsf{sm}(1, \mathbf{u}) & = & \mathbf{u}.
\end{array}
$$

The following notational conventions are used in an arbitrary linear space:

$$
\begin{array}{rcll}
\mathbf{u} + \mathbf{v} & := & \mathsf{add}(\mathbf{u}, \mathbf{v}) & \text{when } \mathbf{u}, \mathbf{v} \in \mathcal{V}, \\
\xi\mathbf{u} & := & \mathsf{sm}(\xi, \mathbf{u}) & \text{when } \xi \in \mathbb{R},\ \mathbf{u} \in \mathcal{V}, \\
-\mathbf{u} & := & \mathsf{opp}(\mathbf{u}) & \text{when } \mathbf{u} \in \mathcal{V}, \text{ and} \\
\mathbf{u} - \mathbf{v} & := & \mathbf{u} + (-\mathbf{v}) & \text{when } \mathbf{u}, \mathbf{v} \in \mathcal{V}.
\end{array}
$$

With this notation, the axioms for a linear space become

$$
\begin{aligned}
(A_1) &\quad \mathbf{u}+(\mathbf{v}+\mathbf{w}) = (\mathbf{u}+\mathbf{v})+\mathbf{w},\\
(A_2) &\quad \mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u},\\
(A_3) &\quad \mathbf{u}+\mathbf{0} = \mathbf{u},\\
(A_4) &\quad \mathbf{u}-\mathbf{u} = \mathbf{0},
\end{aligned}
$$

$$
\begin{aligned}
(S_1) &\quad \xi(\eta\mathbf{u}) = (\xi\eta)\mathbf{u},\\
(S_2) &\quad (\xi+\eta)\mathbf{u} = \xi\mathbf{u}+\eta\mathbf{u},\\
(S_3) &\quad \xi(\mathbf{u}+\mathbf{v}) = \xi\mathbf{u}+\xi\mathbf{v},\\
(S_4) &\quad 1\mathbf{u} = \mathbf{u},
\end{aligned}
$$

valid for all $\xi, \eta \in \mathbb{R}$ and $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathcal{V}$.

Notation: If $\alpha \in \mathbb{R}$, $\mathcal{S}, \mathcal{T} \subset \mathcal{V}$, $\Lambda \subset \mathbb{R}$, and $\mathbf{u} \in \mathcal{V}$, we define

$$
\begin{aligned}
\alpha\mathcal{S} &:= \{\alpha\mathbf{v} \,|\, \mathbf{v} \in \mathcal{S}\},\\
\mathcal{S}+\mathcal{T} &:= \{\mathbf{v}+\mathbf{w} \,|\, \mathbf{v} \in \mathcal{S},\ \mathbf{w} \in \mathcal{T}\},\\
\Lambda\mathcal{S} &:= \{\lambda\mathbf{v} \,|\, \lambda \in \Lambda,\ \mathbf{v} \in \mathcal{S}\},\\
\Lambda\mathbf{u} &:= \Lambda\{\mathbf{u}\}.
\end{aligned}
$$

—**Examples**

1. If S is any set and $\mathcal{V}$ is a linear space, then the set of all mappings from S into $\mathcal{V}$, denoted by $\mathsf{Map}(S, \mathcal{V})$, acquires the structure of a linear space when the operations in $\mathsf{Map}(S, \mathcal{V})$ are defined by value-wise applications of the operations in $\mathcal{V}$. More explicitly, if $f, g \in \mathsf{Map}(S, \mathcal{V})$ and $\alpha \in \mathbb{R}$, then 0, $-f$, $f+g$ and αf are defined by

$$
\begin{aligned}
0(s) &:= 0,\\
(-f)(s) &:= -(f(s)),\\
(f+g)(s) &:= f(s)+g(s), \text{ and}\\
(\alpha f)(s) &:= \alpha f(s)
\end{aligned}
$$

for all $s \in S$.

2. If $\mathcal{V}$ is a linear space and I is a set, then the set $\mathcal{V}^I$ of all families indexed on I with terms in $\mathcal{V}$ acquires the structure of a linear space when the operations in $\mathcal{V}^I$ are defined by term-wise applications of the operations in $\mathcal{V}$. The zero, opposition, addition, and scalar multiplication in $\mathcal{V}^I$ are given as follows for $\mathbf{u}, \mathbf{v} \in \mathcal{V}^I$ and $\alpha \in \mathbb{R}$:

$$\begin{aligned} \mathbf{0}_i &:= \mathbf{0}, \\ (-\mathbf{u})_i &:= -(\mathbf{u}_i), \\ (\mathbf{u}+\mathbf{v})_i &:= \mathbf{u}_i + \mathbf{v}_i, \text{ and} \\ (\alpha\mathbf{u})_i &:= \alpha\mathbf{u}_i \end{aligned}$$

for all $i \in I$. In particular, $\mathcal{V} := \mathbb{R}$ and $I := 1..n$ results in the linear space $\mathbb{R}^n$ of lists of length $n \in \mathbb{N}^\times$; $\mathcal{V} := \mathbb{R}$ and $I := (1..n) \times (1..m)$ gives the linear space $\mathbb{R}^{n\times m}$ of $n \times m$ matrices when $n, m \in \mathbb{N}^\times$.

—Subspaces

Let $\mathcal{V}$ be a linear space.

D02 Definition: *A nonempty subset $\mathcal{U}$ of $\mathcal{V}$ is called a* **subspace of** $\mathcal{V}$ *if it is stable under addition and scalar multiplication; i.e., if $\mathcal{U} + \mathcal{U} \subset \mathcal{U}$ and $\mathbb{R}\mathcal{U} \subset \mathcal{U}$. In this case, $\mathcal{U}$ itself may be considered to be a linear space in a natural way.*

D03 Proposition: *Given a subset $\mathcal{S}$ of $\mathcal{V}$, there is a smallest subspace (with respect to inclusion) which includes $\mathcal{S}$ and which is included in every subspace which includes $\mathcal{S}$. This smallest subspace is called the* **linear span of** $\mathcal{S}$ *and is denoted by* $\mathsf{Lsp}\,\mathcal{S}$.

D04 Definition: *We say that $\mathcal{V}$ is* **finite-dimensional** *if $\mathcal{V} = \mathsf{Lsp}\,\mathcal{S}$ for some finite subset $\mathcal{S}$ of $\mathcal{V}$. The least of the cardinal numbers of such a set $\mathcal{S}$ is called the* **dimension of** $\mathcal{V}$ *and is denoted by* $\dim \mathcal{V}$.

Remark: If $\mathcal{V}$ is finite-dimensional and $\mathcal{U}$ is a subspace of $\mathcal{V}$, then $\mathcal{U}$ is also finite-dimensional.

Assume now that $\mathcal{V}$ is finite-dimensional.

D05 Definition: *Let a finite subset $\mathcal{S}$ of $\mathcal{V}$ be given. We say that $\mathcal{S}$ is* **linearly independent** *if for all $\gamma \in \mathbb{R}^{\mathcal{S}}$,*

$$\sum_{\mathbf{v}\in\mathcal{S}} \gamma_{\mathbf{v}}\mathbf{v} = 0 \implies (\gamma_{\mathbf{v}} = 0 \text{ for all } \mathbf{v} \in \mathcal{S}).$$

We say that $\mathcal{S}$ is **linearly dependent** *if $\mathcal{S}$ is not linearly independent.*

If $\mathcal{U}$ is a subspace of $\mathcal{V}$, we say that $\mathcal{S}$ **spans** *$\mathcal{U}$ if $\mathsf{Lsp}\ \mathcal{S} = \mathcal{U}$.*

If $\mathcal{S}$ is linearly independent and spans $\mathcal{V}$, we say that $\mathcal{S}$ is a **basis of** *$\mathcal{V}$.*

If $m \in \mathbb{N}^{\times}$ and $\mathbf{b} = (\mathbf{b}_i \,|\, i \in 1..m) \in \mathcal{V}^m$ is such that

$$\mathbf{b}_i = \mathbf{b}_j \implies i = j$$

for all $i, j \in 1..m$ and $\{\mathbf{b}_i \,|\, i \in 1..m\}$ is a basis of $\mathcal{V}$, we say that $\mathbf{b}$ is a **list-basis of** *$\mathcal{V}$.*

D06 Proposition: *Let $\mathcal{S} \subset \mathcal{V}$ be given, and put $n := \mathsf{dim}\ \mathcal{V}$. If $\mathcal{S}$ is linearly independent, then $\#\mathcal{S} \le n$. If $\mathcal{S}$ spans $\mathcal{V}$, then $\mathcal{S}$ includes a linearly independent set of n members. If $\mathcal{S}$ is a basis of $\mathcal{V}$, then $\#\mathcal{S} = n$.*

D07 Definition: *We say that a given pair $(\mathcal{U}_1,\mathcal{U}_2)$ of subspaces of $\mathcal{V}$ is* **supplementary (in $\mathcal{V}$)** *if:*

$$\begin{array}{ll} (i) & \mathcal{U}_1 + \mathcal{U}_2 = \mathcal{V}, \text{ and} \\ (ii) & \mathcal{U}_1 \cap \mathcal{U}_2 = \{\mathbf{0}\}. \end{array}$$

In this case, we say that $(\mathcal{U}_1,\mathcal{U}_2)$ is a **decomposition of** *$\mathcal{V}$.*

D08 Proposition: *Let $\mathcal{U}_1$ and $\mathcal{U}_2$ be subspaces of $\mathcal{V}$ such that $\mathcal{U}_1 \cap \mathcal{U}_2 = \{\mathbf{0}\}$. Then*

$$\mathsf{dim}\ \mathcal{U}_1 + \mathsf{dim}\ \mathcal{U}_2 \le \mathsf{dim}\ \mathcal{V}.$$

Equality is obtained if and only if $\mathcal{U}_1 + \mathcal{U}_2 = \mathcal{V}$; i.e., if and only if $\mathcal{U}_1$ and $\mathcal{U}_2$ are supplementary.

D09 Proposition: *If* $\mathcal{U}_1$ *and* $\mathcal{U}_2$ *are subspaces of* $\mathcal{V}$*, then any two of the following statements imply the third:*

$$\begin{aligned} &(i) && \mathcal{U}_1 \cap \mathcal{U}_2 = \{\mathbf{0}\}, \\ &(ii) && \mathcal{U}_1 + \mathcal{U}_2 = \mathcal{V}, \\ &(iii) && \mathsf{dim}\, \mathcal{U}_1 + \mathsf{dim}\, \mathcal{U}_2 = \mathsf{dim}\, \mathcal{V}. \end{aligned}$$

—Linear Mappings

Let $\mathcal{V}$ and $\mathcal{W}$ be finite-dimensional linear spaces.

D10 Definition: *A mapping* $\mathbf{L} : \mathcal{V} \to \mathcal{W}$ *is said to be* **linear** *if it preserves addition and scalar multiplication; i.e., if*

$$\mathbf{L}(\mathbf{v}_1 + \mathbf{v}_2) = \mathbf{L}\mathbf{v}_1 + \mathbf{L}\mathbf{v}_2$$

for all $\mathbf{v}_1, \mathbf{v}_2 \in \mathcal{V}$ *and*

$$\mathbf{L}(\alpha\mathbf{v}) = \alpha\mathbf{L}\mathbf{v}$$

for all $\alpha \in \mathbb{R}$ *and* $\mathbf{v} \in \mathcal{V}$.

Let a linear mapping $\mathbf{L} : \mathcal{V} \to \mathcal{W}$ be given.

D11 Definition: *We define the* **null space of L** *and the* **range of L** *by*

$$\mathsf{Null}\, \mathbf{L} := \{\mathbf{v} \in \mathcal{V} \,|\, \mathbf{L}\mathbf{v} = \mathbf{0}\}$$

and

$$\mathsf{Rng}\, \mathbf{L} := \{\mathbf{L}\mathbf{v} \,|\, \mathbf{v} \in \mathcal{V}\},$$

respectively.

D12 Proposition: $\mathsf{Null}\, \mathbf{L}$ *is a subspace of* $\mathcal{V}$ *and* $\mathsf{Rng}\, \mathbf{L}$ *is a subspace of* $\mathcal{W}$.

D13 Proposition: $\mathbf{L}$ *is injective if and only if* $\mathsf{Null}\, \mathbf{L} = \{\mathbf{0}\}$.

D14 Proposition: *We have*

$$\mathsf{dim}\, \mathsf{Null}\, \mathbf{L} + \mathsf{dim}\, \mathsf{Rng}\, \mathbf{L} = \mathsf{dim}\, \mathcal{V}.$$

D15 Proposition: (Linear Pigeonhole Principle): *If* $\mathbf{L}$ *is injective [surjective], then* $\mathsf{dim}\ \mathcal{V} \leq \mathsf{dim}\ \mathcal{W}$ $[\mathsf{dim}\ \mathcal{V} \geq \mathsf{dim}\ \mathcal{W}]$. *Equality holds in either case if and only if* $\mathbf{L}$ *is invertible.*

D16 Proposition: *If* $(\mathcal{U}_1, \mathcal{U}_2)$ *is a decomposition of* $\mathcal{V}$, *then there is exactly one linear mapping* $\mathbf{P} : \mathcal{V} \to \mathcal{U}_1$ *such that* $\mathbf{Pu} = \mathbf{u}$ *for all* $\mathbf{u} \in \mathcal{U}_1$ *and* $\mathbf{Pu} = \mathbf{0}$ *for all* $\mathbf{u} \in \mathcal{U}_2$. $\mathbf{P}$ *is called the* **projection of** $\mathcal{V}$ **onto** $\mathcal{U}_1$ **along** $\mathcal{U}_2$.

D17 Definition: *The set of all linear mappings from* $\mathcal{V}$ *to* $\mathcal{W}$ *is denoted by* $\mathsf{Lin}(\mathcal{V}, \mathcal{W})$. *We abbreviate* $\mathsf{Lin}\ \mathcal{V} := \mathsf{Lin}(\mathcal{V}, \mathcal{V})$. *Members of* $\mathsf{Lin}\ \mathcal{V}$ *are called* **lineons**.

D18 Definition: *Let* $\mathbf{L} \in \mathsf{Lin}\ \mathcal{V}$ *be given. A subspace* $\mathcal{U}$ *of* $\mathcal{V}$ *is said to be an* **L-space** *if* $\mathbf{L}_{>}(\mathcal{U}) \subset \mathcal{U}$. *If* $\mathcal{U}$ *is an* $\mathbf{L}$*-space, we define the mapping* $\mathbf{L}_{|\mathcal{U}} : \mathcal{U} \to \mathcal{U}$ *by* $\mathbf{L}_{|\mathcal{U}}(\mathbf{u}) := \mathbf{Lu}$ *for all* $\mathbf{u} \in \mathcal{U}$.

D19 Proposition: *Assume that* $\mathsf{dim}\ \mathcal{V} \geq 2$. *Then every lineon in* $\mathsf{Lin}\ \mathcal{V}$ *admits at least one two-dimensional* $\mathbf{L}$*-space.*

Remark: A proof of this Proposition is non-trivial and depends on the Fundamental Theorem of Algebra (see, for example, §**94** of [**7**]).

D20 Proposition: *Put* $n := \mathsf{dim}\ \mathcal{V}$. *Let* $\mathbf{L} \in \mathsf{Lin}\ \mathcal{V}$ *and a list-basis* $\mathbf{b} = (\mathbf{b}_i \,|\, i \in 1..n)$ *of* $\mathcal{V}$ *be given. Then there is exactly one matrix in* $\mathbb{R}^{n \times n}$, *denoted by* $[\mathbf{L}]_{\mathbf{b}}$, *such that*

$$\mathbf{Lb}_i = \sum_{j \in 1..n} ([\mathbf{L}]_{\mathbf{b}})_{ij}\, \mathbf{b}_j$$

for all $i \in 1..n$. $[\mathbf{L}]_{\mathbf{b}}$ *is called the* **matrix of L relative to b**.

Note that if $\mathsf{dim}\ \mathcal{V} = 4$, then for every list-basis $\mathbf{b}$ of $\mathcal{V}$, we have

$$[\mathbf{1}_{\mathcal{V}}]_{\mathbf{b}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

D21 Proposition: *Put* $n := \mathsf{dim}\ \mathcal{V}$, *and let* $\mathbf{L} \in \mathsf{Lin}\ \mathcal{V}$ *be given. Then*

$$\sum_{i \in 1..n} ([\mathbf{L}]_{\mathbf{b}})_{ii}$$

is independent of the choice of list-basis $\mathbf{b}$ *of* $\mathcal{V}$. *This common value is called the* **trace of L**.

—**Tensor Products**

Assume that $\mathcal{V}$ is an inner-product space (see §**5.1**).

D22 Definition: *Let* $\mathbf{a}, \mathbf{b} \in \mathcal{V}$ *be given. The* **tensor product of a and b**, *denoted by* $\mathbf{a} \otimes \mathbf{b} : \mathcal{V} \to \mathcal{V}$, *is defined by*

$$(\mathbf{a} \otimes \mathbf{b})(\mathbf{v}) := (\mathbf{b} \cdot \mathbf{v})\mathbf{a} \quad \textit{for all } \mathbf{v} \in \mathcal{V}.$$

D23 Proposition: *For all* $\mathbf{a}, \mathbf{b} \in \mathcal{V}$, $\mathbf{a} \otimes \mathbf{b}$ *is linear; i.e.,* $\mathbf{a} \otimes \mathbf{b} \in \mathsf{Lin}\ \mathcal{V}$.

References

1. Einstein, Albert, *Relativity,* Bonanza Books, New York, 1961.

2. Einstein, Albert, H. A. Lorentz, H. Minkowski, and H. Weyl, *The Principle of Relativity,* Dover Publications, Inc., Toronto, 1952.

3. Minkowski, H., *Raum und Zeit,* lecture given in 1908.

4. Newton, Isaac, *Mathematical Principles of Natural Philosophy,* Chicago: Encyclopædia Britannica, Inc., 1952.

5. Noll, Walter, "Euclidean geometry and Minkowskian chronometry", *American Mathematical Monthly,* **71**, 129-144 (1964).

6. Noll, Walter, "Space-time structures in classical mechanics", *Delaware Seminar in the Foundations of Physics,* pp. 28-34, Springer-Verlag, New York, (1967).

7. Noll, Walter, *Finite-Dimensional Spaces,* Kluwer, Boston, 1987.

8. Noll, Walter, and Schäffer, Juan Jorge, "Order-Isomorphisms in Affine Spaces", *Annali di Matematica pura ed applicata*, (4) 117, 243-262 (1978).

9. Rudin, Walter, *Real and Complex Analysis,* Mc-Graw Hill Book Co., New York, 1987.

10. Schutz, John, *Foundations of Special Relativity: Kinematic Axioms for Minkowski Space-Time,* Springer-Verlag, New York, 1973.

The following two books, although not referenced, provided many ideas useful in the writing of the text.

Naber, Gregory, *The Geometry of Minkowski Spacetime: An Introduction to the Mathematics of the Special Theory of Relativity,* Springer-Verlag, New York, 1992.

Rindler, Wolfgang, *Essential Relativity,* Springer-Verlag, New York, 1977.

Index