Solutions for Exercises

contained in the chapter

Albrecht Iseler:
Über richtungsbasierte Relationen in reellen Vektorräumen

in

K. Ch. Klauer & H. Westmeyer:
Psychologische Methoden und soziale Prozesse
Festschrift für Hubert Feger
Lengerich: 1998,
pp. 80-121

The subsequent solutions are given in English to support quotations of results in other papers. For the same reason, we will sometimes refer to the English definitions in the appendix instead of the German ones in the main text.

The above quoted chapter will subsequently be called the main text.

Before solutions are tried, the following corrections to the main text should be noted:

- In the 6th line of p. 93, the exponent in the definition of α_i must be $i-1$ instead of $2-i$; i.e., $\alpha_i := \lambda^{i-1} / (1+\lambda)$.
- In Definition A.2.g, the subscripts are displaced. The first equation must be $g(\mathbf{x}_1 + \mathbf{x}_2) = g(\mathbf{x}_1) + g(\mathbf{x}_2)$.

Links to additional documents (an extended version of the appendix of the main text, including proofs, and further errata) are contained in the WWW-document http://userpage.fu-berlin.de/~iseler/papers/dirbas_mat.htm.

To prevent confusions between the numberings of propositions, equations etc., the numbering x.y resp. (x.y) is subsequently chosen such that x is the number of the respective exercise plus 100. Other numbers of equalities, definitions, propositions etc. refer to the main text, and the same holds for numbers of pages in the range 80-121.

We resume the general notational conventions from pp. 81−82 of the main text. In particular, recall that propositions etc. containing expressions with subscript i or j refer to values 1 and 2 of these subscripts unless another range is specified in the local context.

The author is aware that the solution of the exercises may be for some readers their first experience with general notions of real vector spaces. Since these readers may be uncertain about the applicability of wellknown properties of $\mathbb{R}^n$, such properties will sometimes be derived from requirements of Definitions A.1 and A.2 in the main text. However, a consistent application of this approach would make the solutions tedious and hard to follow up. For this reason, the reader is encouraged to try solutions under the simplifying assumption that the underlying vector space is an $\mathbb{R}^n$. A basic familiarity with this class of vector spaces should be sufficient for solutions under the said simplifying assumption. An additional hint refers to Exercise 13. Whereas a general solution of this exercise requires some advanced analytical tools, it becomes elementary under the simlifying assumption that S is the power set of a finite set W.

Subsequently, some exercises will be solved in a way yielding as side results some generalisations, whose discussion in the main text would have diverted from its train of thoughts. Naturally, many readers will choose other solutions.

Exercise 1, p. 87

Let a binary relation R on a real vector space F be based on a linear map g:$F\to\mathbb{R}$ such that the defining equivalence

$$\mathbf{x}_1 \, R \, \mathbf{x}_2 \;:\Leftrightarrow\; g(\mathbf{x}_1) \le g(\mathbf{x}_2) \qquad (101.1)$$

holds for all elements $\mathbf{x}_1$ and $\mathbf{x}_2$ of F, and consider the set C_R defined by

$$C_R := \{\mathbf{x} \in F\text{: } g(\mathbf{x}) \ge 0\}. \qquad (101.2)$$

Certainly, this set is a cone (i.e., closed under multiplication by scalars $\lambda > 0$): For every scalar $\lambda > 0$ and $\mathbf{x} \in C_R$, the linearity of the map g implies

$$g(\lambda\, \mathbf{x}) = \lambda\, g(\mathbf{x}) \ge 0, \qquad (101.3)$$

the last inequality following from the assumptions $\lambda > 0$ and $\mathbf{x} \in C_R$ by the definition of the set C_R in Equation (101.2). The same definition allows to derive $\lambda\, \mathbf{x} \in C_R$ from $\lambda\, \mathbf{x} \ge 0$.

To verify that the Equivalence (1.1) holds for all elements $\mathbf{x}_1$ and $\mathbf{x}_2$ of F, we can combine Equivalence (101.1) and Equation (101.2) to the chain

$\mathbf{x}_1\ R\ \mathbf{x}_2 \ \Leftrightarrow\ g(\mathbf{x}_1) \leq g(\mathbf{x}_2) \ \Leftrightarrow\ g(\mathbf{x}_2 - \mathbf{x}_1) \geq 0 \Leftrightarrow\ \mathbf{x}_2 - \mathbf{x}_1 \in C_R,$ (101.4)

the second equivalence in this chain being based on the linearity of the map g, which implies $g(\mathbf{x}_2 - \mathbf{x}_1) = g(\mathbf{x}_2) - g(\mathbf{x}_1)$.

Exercise 2, p. 87

The cone C_R defined by Equation (101.2) is pointed, since the linearity of the map g implies $g(\mathbf{0}) = 0$. To verify its convexity, let $\mathbf{x}_1$ and $\mathbf{x}_2$ be arbitrary elements of C_R and consider the vector $\mathbf{x}$ defined by

$\mathbf{x} := \alpha\ \mathbf{x}_1 + (1-\alpha)\ \mathbf{x}_2$ (102.1)

with $\alpha \in [0, 1]$. We have to show that every such $\mathbf{x}$ is an element of C_R. The linearity of g implies

$g(\mathbf{x}) = \alpha\ g(\mathbf{x}_1) + (1-\alpha)\ g(\mathbf{x}_2) \geq 0,$ (102.2)

the final inequality following from the assumptions $\alpha \in [0, 1]$ and $\mathbf{x}_\mathrm{i} \in C_R$, the second one leading to $g(\mathbf{x}_\mathrm{i}) \geq 0$ by Equation (101.2).

Since linear orders have been introduced (on p. 83) as binary relations R one a real vector space F, where Equivalence (1.1) holds for all elements $\mathbf{x}_1$ and $\mathbf{x}_2$ of F and a pointed, convex cone C_R, we may combine the results of Exercises 1 and 2 in the conclusion that R is a linear order, indeed.

Exercise 3, p. 87

In the situation of the preceeding exercises, let $\mathbf{x}$ and $\mathbf{y}$ be elements of F such that the relation $\mathrm{n}\ \mathbf{x}\ R\ \mathbf{y}$ holds for every $\mathrm{n} \in \mathbb{N}$. Then Equivalence (101.1) leads to $\mathrm{n}\ g(\mathbf{x}) \leq g(\mathbf{y})$ for every $\mathrm{n} \in \mathbb{N}$. (Note that the linearity of the map g implies $g(\mathrm{n}\ \mathbf{x}) = \mathrm{n}\ g(\mathbf{x})$.) But then we get $g(\mathbf{x}) \leq 0 = g(\mathbf{0})$, and another reference to Equivalence (101.1) allows to derive $\mathbf{x}\ R\ \mathbf{0}$.

Exercise 4, p. 88

To be proved: If R is a binary relation on a subset Y of e real vector space F, and C_R a cone in F such that the eqivalence

$\mathbf{y}_1\ R\ \mathbf{y}_2 \ \Leftrightarrow\ \mathbf{y}_2 - \mathbf{y}_1 \in C_R$ (104.1)

holds for all elements $\mathbf{y}_1$ and $\mathbf{y}_2$ of Y, then the equivalence

$\mathbf{y}_{11}\ R\ \mathbf{y}_{12} \ \Leftrightarrow\ \mathbf{y}_{21}\ R\ \mathbf{y}_{22}$ (104.2) must hold for all elements $\mathbf{y}_{ij}$ of Y, where the equation

$$\mathbf{y}_{22}-\mathbf{y}_{21} = \lambda\,(\mathbf{y}_{12}-\mathbf{y}_{11}) \qquad (104.3)$$

holds for some scalar $\lambda > 0$.

So let Y, R and C_R with these properties be given. For elements $\mathbf{y}_{ij}$ of Y fulfilling Equation (104.3) for a scalar $\lambda > 0$, the outer equivalences in the chain

$$\mathbf{y}_{11}\,R\,\mathbf{y}_{12} \;\Leftrightarrow\; \mathbf{y}_{12}-\mathbf{y}_{11} \in C_R \;\Leftrightarrow\; \lambda\,(\mathbf{y}_{12}-\mathbf{y}_{11}) \in C_R \;\Leftrightarrow\; \mathbf{y}_{21}\,R\,\mathbf{y}_{22} \qquad (104.4)$$

follow from an application of Equivalence (104.1), whereas the second equivalence in the chain is granted by the assumption that C_R is a cone.

Exercise 5, p. 89

We have to prove Corollary 2.2. So let R be a binary relation on a real vector space F. Since the implication (3)⇒(1) has already been proved in Exercise 4 (without explicit reference to the concept of a direction-based relation), it suffices to verify the implications (1)⇒(2)⇒(3) and the concluding assertion.

(1)⇒(2): Assume that the relation R is direction based, let $\mathbf{y}_{11}$, $\mathbf{y}_{12}$, and $\mathbf{y}$ be arbitrary elements of F, and $\lambda > 0$ a scalar. With the definition

$$\mathbf{y}_{2j} := \lambda\,\mathbf{y}_{1j} \qquad (105.1)$$

for j = 1 and j = 2, we can easily verify the validity of Equation (104.3); hence Equivalence (104.2) follows from the assumption that R is direction-based. But as special cases of this result for $\lambda = 1$ resp. $\mathbf{y} = \mathbf{0}$, we obtain that the relation R is, indeed, invariant under translation (resp. under multiplication by strictly positive scalars) .

(2)⇒(3): Under the assumption that the relation R is invariant under translation and multiplication by strictly positive scalars, define a subset C_R of F by

$$C_R := \{\mathbf{x} \in F\colon \mathbf{0}\,R\,\mathbf{x}\}. \qquad (105.2)$$

This definition implies the outer equivalences in the following chain, whereas the second one follows for every $\lambda > 0$ from the assumption that the relation R is invariant under multiplication by strictly positive scalars:

$$\mathbf{x} \in C_R \;\Leftrightarrow\; \mathbf{0}\,R\,\mathbf{x} \;\Leftrightarrow\; \mathbf{0}\,R\,\lambda\,\mathbf{x} \;\Leftrightarrow\; \lambda\,\mathbf{x} \in C_R \qquad (105.3)$$

But if this equivalence holds for every $\mathbf{x} \in C_R$ and every scalar $\lambda > 0$, then C_R must be a cone (in the weak understanding of this concept explicated in Definition A.2.c). To establish the validity of Equivalence (1.1) for all elements $\mathbf{x}_1$ and $\mathbf{x}_2$ of F, observe that the first equivalence in the follwoing chain is granted by the assumption that the relation R is invariant under translation (e.g., by $-\mathbf{x}_1$), whereas the second equivalence results from Equation (105.2).

$$\mathbf{x}_1 \, R \, \mathbf{x}_2 \;\Leftrightarrow\; \mathbf{0} \, R \, \mathbf{x}_2 - \mathbf{x}_1 \;\Leftrightarrow\; \mathbf{x}_2 - \mathbf{x}_1 \in C_R \qquad (105.4)$$

Finally, observe that the concluding assertion of Corollary 2.2 is a side result of the above proof of the implication (2)⇒(3). In particular, to fulfill Equivalence (1.1) for $\mathbf{x}_1 = \mathbf{0}$, a set C_R must be identical with the one defined in Equation (105.2).

To round the result, note that Corollary 2.2 shows the equivalence of three possible definitions of linear orders: Adding one of the equivalent properties desribed in Assertions (1), (2) and (3) to the requirement that R is reflexive and transitive, we get a linear order. In the solution to Ercercise 10, a fourth equivalent approach will be added,

Exercise 6, p. 89

Let Y be the set consisting of the vectors $\mathbf{y}_{ij}$ in the right part of Figure 1. We have to prove the following assertion: There exists a position of the zero-element and a binary relation R on Y with the two invariances claimed in Assertion (2) of Corollary 2.2 such that $\mathbf{y}_{11} \, R \, \mathbf{y}_{12}$ and $\neg(\mathbf{y}_{21} \, R \, \mathbf{y}_{22})$. Now consider a situation with

$$\mathbf{y}_{11} = \mathbf{0}, \qquad (106.1)$$

and

$$R = \{(\mathbf{y}_{11}, \mathbf{y}_{12})\}. \qquad (106.2)$$

I.e., the zero-element of tthe vector space is identical with $\mathbf{y}_{11}$, and the ordered pair $(\mathbf{y}_{11}, \mathbf{y}_{12})$ is the only ordered pair $(\mathbf{y}', \mathbf{y}'')$ of elements of Y where the relation $\mathbf{y}' \, R \, \mathbf{y}''$ holds.

Since Equation (106.2) implies $\mathbf{y}_{11} \, R \, \mathbf{y}_{12}$ as well as $\neg(\mathbf{y}_{21} \, R \, \mathbf{y}_{22})$, it suffices to verify that this situation is compatible with the invariances under consideration. Now these invariances refer to situations, where $\mathbf{y}'_{ij}$ are (not necessarily different) elements of Y. For such situations, the equivalence

$$\mathbf{y}'_{11} \, R \, \mathbf{y}'_{12} \;\Leftrightarrow\; \mathbf{y}'_{21} \, R \, \mathbf{y}'_{22} \qquad (106.3)$$

is claimed under two conditions: It must hold

- If there is an element $\mathbf{x}$ of F such that

$$\mathbf{y}'_{2j} = \mathbf{y}'_{1j} + \mathbf{x}, \qquad (106.4)$$

- If there is a scalar $\lambda > 0$ such that

$\mathbf{y'}_{2j} = \lambda\, \mathbf{y'}_{1j}$, (106.5)

the validity of Equation (106.4) resp. (106.5) being required for j = 1 and j = 2 with the same **x** resp. with the same λ. Of course, Equivalence (106.3) is tautologically true for $\mathbf{x} = \mathbf{0}$ resp. $\lambda = 1$: Then $\mathbf{y}_{2j} = \mathbf{y}_{1j}$ follows immediately. There is also a non-tautological consequence of invariance under translation. Consider a situation with $\mathbf{y'}_{11} = \mathbf{y'}_{12}$ and $\mathbf{y'}_{21} = \mathbf{y'}_{22}$. Here Equation (106.4) holds for $\mathbf{x} = \mathbf{y'}_{21} - \mathbf{y'}_{11}$. Requiring the validity of Equivalence (106.3) for such situations means that the relation *R* must be either consistently reflexive or consistently non-reflexive. Indeed, the relation specified by Equation (106.2) is consistently non-reflexive. But for $\mathbf{y'}_{11} \neq \mathbf{y'}_{12}$, there isn't any $\mathbf{x} \neq \mathbf{0}$ such that Equation (106.4) holds for elements $\mathbf{y'}_{ij}$ of *Y*. (Try all triples of elements $\mathbf{y'}_{11}$, $\mathbf{y'}_{12}$ and $\mathbf{y'}_{21}$ of *Y* with $\mathbf{y'}_{11} \neq \mathbf{y'}_{12}$ and $\mathbf{y'}_{11} \neq \mathbf{y'}_{21}$, and verify that the vector $\mathbf{y'}_{12} + (\mathbf{y'}_{21} - \mathbf{y'}_{11})$ is not contained in *Y*. So there is no requirement of invariance under translation beyond consistent relexivity or non-reflexivity.

For invariance under multiplication by strictly positive scalars, consider the following geometric interpretation of Equation (106.5) with $0 < \lambda < 1$. It means that $\mathbf{y'}_{2j}$ is an element of the line segment connecting the zero-element and $\mathbf{y'}_{1j}$. Similarly, for $\lambda > 1$ the equation implies that $\mathbf{y'}_{1j}$ is an element of the line segment connecting the zero-element and $\mathbf{y'}_{2j}$. But it is obvious that no configuration of this kind can be made from the points $\mathbf{y}_{ij}$ in the left part of Figure 1, if the zero-element is identical with $\mathbf{y}_{11}$ (Equation (106.1)). So the invariance under multiplication by strictly positive scalars cannot be violated by the situation specified by Equations (106.1) and (106.2).

The above geometric approach can also help to specify the only position of the zero-element, where invariance under multiplication by strictly positive scalars has non-trivial consequences: Let **y** be the intersection point of the two lines containing the points $\mathbf{y}_{11}$ and $\mathbf{y}_{21}$ resp. $\mathbf{y}_{12}$ and $\mathbf{y}_{22}$, and consider a situation where $\mathbf{y} = \mathbf{0}$. Geometric intuition suggests that in this case the equation

$\mathbf{y}_{2j} = 1.5\, \mathbf{y}_{1j}$ (106.6)

must hold in a situation with the given property

$\mathbf{y}_{22} - \mathbf{y}_{21} = 1.5\, (\mathbf{y}_{12} - \mathbf{y}_{11})$. (106.7)

To verify this intuition algebraically, we can easily derive the equation

$\mathbf{y}_{11} + 2\, (\mathbf{y}_{11} - \mathbf{y}_{21}) = \mathbf{y}_{12} + 2\, (\mathbf{y}_{12} - \mathbf{y}_{22})$. (106.8)

Equation (106.7) allows to replace $\mathbf{y}_{21}$ on the left hand side of Equation (106.8) by $\mathbf{y}_{22} - 1.5\, (\mathbf{y}_{12} - \mathbf{y}_{11})$, and then a simple rearrangement of terms gives the right hand side of Equation (106.8). Under a geometric interpretation, the point $\mathbf{y}_{11} + 2\, (\mathbf{y}_{11} - \mathbf{y}_{21})$ is the result of the following operation: Extend the line connecting the points $\mathbf{y}_{11}$ and **y**21 beyond $\mathbf{y}_{11}$, and - starting at $\mathbf{y}_{11}$ - step along this extension twice the difference $\mathbf{y}_{11}-\mathbf{y}_{21}$. An application of the same interpretation to the right hand side of Equation (106.8), this equation shows that both operations result in the same point; i.e., the resulting point belongs to both extensions. So it is the intersection of the two extensions, for which we have introduced the denotation $\mathbf{y}$. Then Equation (106.6) can be derived, indeed, from the assumption $\mathbf{y} = \mathbf{0}$: If both sides of Equation (106.8) equal $\mathbf{0}$, we can write $\mathbf{0} = 3\ \mathbf{y}_{1j} - 2\ \mathbf{y}_{2j}$, and then a rearrangement of terms leads to Equation (106.6) . But then the joint validity of $\mathbf{y}_{11}\ R\ \mathbf{y}_{12}$ and $\neg(\mathbf{y}_{21}\ R\ \mathbf{y}_{22})$ would be excluded by invariance under multiplication by strictly positive scalars.

Although the existence of just one position of the zero-element and one relation R with the required properties is sufficient, it may be interesting to give a more general description of requirements following from invariance under translation and under multiplication by strictly positive scalar:

- Independendent of the position of the zero-element, invariance under translation implies that a binary relation R on our set Y must be consistently reflexive or consistently non-reflexive.
- The equivalence $\mathbf{y}_{11}\ R\ \mathbf{y}_{12} \Leftrightarrow \mathbf{y}_{21}\ R\ \mathbf{y}_{22}$ is necessary, if and only if the equation $\mathbf{y}_{1j} + 2\ (\mathbf{y}_{1j} - \mathbf{y}_{2j})$ holds; i.e., iff $\mathbf{y} = \mathbf{0}$.

Every binary relation R on Y fulfilling these requirements is invariant under translation and multiplication by strictly positive scalars.

The precise meaning of the invariances may also be clarified by to erroneous approaches for situations with $\mathbf{y} \neq \mathbf{0}$. One could be tempted to define points $\mathbf{y}'_{ij} := \mathbf{y}_{ij} - \mathbf{y}$ and consider the chain

$$\mathbf{y}_{11}\ R\ \mathbf{y}_{12}\ \ ?\!\Leftrightarrow\ \ \mathbf{y}'_{11}\ R\ \mathbf{y}'_{12}\ \ ?\!\Leftrightarrow\ \mathbf{y}'_{21}\ R\ \mathbf{y}'_{22}\ \ ?\!\Leftrightarrow\ \ \mathbf{y}_{21}\ R\ \mathbf{y}_{22}$$

of questionable equivalences. Indeed, if the vectors $\mathbf{y}'_{ij}$ would be elements of the set Y, then the outer equivalences would follow from invariance under translation (by $-\mathbf{y}$), whereas the second equivalence would be granted by invariance under multiplication by the strictly positive scalar 1.5. (it may be left to the reader to verify these claims.) However, if the relation R is a relation on a set Y consisting only of the vectors $\mathbf{y}_{ij}$, then the above temptative chain of equivalences would lead us outside the domain of the relation.

A similar argument applies to another consideration: From the above definition of a vector $\mathbf{y}$, we can derive $\mathbf{y}_{2j} = 1.5\ \mathbf{y}_{1j} - 0.5\ \mathbf{y}$. In other words, the transition from $\mathbf{y}_{1j}$ to $\mathbf{y}_{2j}$ is considered as a concatenation of a multiplication by $\lambda = 1.5$ and a translation by $\mathbf{x} = -0.5\mathbf{y}$. But again, this concatenation has consequences only if the results of the first operation (the products $1.5\ \mathbf{y}_{1j}$) are elements of Y, i.e., if $\mathbf{y} = \mathbf{0}$. More generally, such concatenations can very well be used to demonstrate the difference between direction-basedness and invariance under translation and multiplication by strictly positive scalars: Whereas these invariances grant the equivalence

$$\lambda \mathbf{y}_1 + \mathbf{x} \, R \, \lambda \mathbf{y}_2 + \mathbf{x} \;\Leftrightarrow\; \mathbf{y}_1 \, R \, \mathbf{y}_2 \qquad (106.9)$$

for $\mathbf{y}_j \in Y$, $\lambda > 0$ and $\mathbf{x} \in F$ only in situations with $\lambda \mathbf{y}_j \in Y$, this additional requirement is unnecessary for a direction-based relation. (See the equivalence of Assertions (5a) and (5b) in Lemma A.6.) Of course, considerations of this kind are unnecessary, if a relation R is defined on an entire vector space F: For $Y = F$, the condition $\lambda \mathbf{y}_j \in Y$ will always hold.

Exercise 7, p. 93

It has already been pointed out in the remarks preceeding the solututions that the definition of the coefficient α_i is unfortunately wrong in the printed chapter. It must be

$$\alpha_i := \lambda^{i-1} / (1+\lambda), \qquad (107.1)$$

where λ is the scalar underlying the assumption

$$\mathbf{y}_{22} - \mathbf{y}_{21} = \lambda \, (\mathbf{y}_{12} - \mathbf{y}_{11}). \qquad (107.2)$$

Combining Equation (107.1) with the assumption $\lambda > 0$, we get $0 < \alpha_i < 1$. So the two linear combinations $\mathbf{y}_{ij} + \alpha_1 (\mathbf{x}_i - \mathbf{y}_{ij})$ for i = 1 and i = 2 are elements of the line segment from $\mathbf{y}_{ij}$ to $\mathbf{x}_i$. It suffices to verify that both linear combinations result in the same vector: Then this result must be the intersection point. For greater generality, the equality of both points is subsequently established for vectors $\mathbf{x}_i$ fulfilling the equation

$$\mathbf{x}_i = \gamma \, \mathbf{y}_{k1} + (1-\gamma) \, \mathbf{y}_{k2} \qquad (107.3)$$

with k := 3−i and an arbitrary scalar γ, which is 0.3 in the right part of Figure 1. For every such situation, we can write:

$$\begin{aligned}
\mathbf{y}_{1j} + \alpha_1 (\mathbf{x}_1 - \mathbf{y}_{1j}) &= \mathbf{y}_{1j} + (1 / (1+\lambda)) \, (\gamma \, \mathbf{y}_{21} + (1-\gamma) \, \mathbf{y}_{22} - \mathbf{y}_{1j}) \\
&= (\lambda \, \mathbf{y}_{1j} + \mathbf{y}_{22} - \gamma \, (\mathbf{y}_{22} - \mathbf{y}_{21})) / (1 + \lambda) \\
&= (\lambda \, \mathbf{y}_{12} + \mathbf{y}_{2j} - \gamma \, \lambda \, (\mathbf{y}_{12} - \mathbf{y}_{11})) / (1 + \lambda) \\
&= \mathbf{y}_{2j} + (\lambda / (1+\lambda)) \, (\gamma \, \mathbf{y}_{11} + (1-\gamma) \, \mathbf{y}_{12} - \mathbf{y}_{2j}) \\
&= \mathbf{y}_{2j} + \alpha_2 (\mathbf{x}_2 - \mathbf{y}_{2j}). \qquad (107.4)
\end{aligned}$$

The outer equalities are immediate applications of Equations (107.1) and (107.3), whereas the second equality as well as the last but one are based on simple algebraic transformations. For the transition from the second line to the third one, observe first that the changes immediately after γ are justified by Equation (107.2). The further changes are based on the equation

$$\lambda \mathbf{y}_{1j} + \mathbf{y}_{22} = \lambda \mathbf{y}_{12} + \mathbf{y}_{2j}. \tag{107.5}$$

For j = 2, this equality is tautological, whereas its validity for j = 1 follows from a suitable rearrangement of terms in Equation (107.2).

Although Equation (107.4) holds for every scalar γ, its interpretation has to be qualified for $\gamma = 0$ and $\gamma = 1$. For $\gamma = 1$, we get $\mathbf{x}_1 = \mathbf{y}_{21}$ and $\mathbf{x}_2 = \mathbf{y}_{11}$. So the line segments from $\mathbf{y}_{11}$ to $\mathbf{x}_1$ and from $\mathbf{y}_{21}$ to $\mathbf{x}_2$ are identical and don't have a unique intersection point. Nevertheless, we could pick out one point $\mathbf{y'}_1$ from this line segment by the definition $\mathbf{y'}_1 := \mathbf{y}_{i1} + \alpha_i (\mathbf{x}_i - \mathbf{y}_{i1})$ for such situations, the identity of the results for i = 1 and i = 2 being established by Equation (107.4). Then the equivalence

$$\mathbf{y'}_1 \, R \, \mathbf{y'}_2 \Leftrightarrow \mathbf{y}_{i1} \, R \, \mathbf{y}_{i2} \tag{107.6}$$

holds again for $i \in \{1, 2\}$, and Equivalence (2.2) follows.

Similarly for $\gamma = 0$: Then the line segment from $\mathbf{y}_{12}$ (= $\mathbf{x}_2$) to $\mathbf{y}_{22}$ (= $\mathbf{x}_1$) contains the point given by the definition $\mathbf{y'}_2 := \mathbf{y}_{i2} + \alpha_i (\mathbf{x}_i - \mathbf{y}_{i2})$.

Exercise 8, p. 93

For $F := \mathbb{R}^2$, let the subset Y of F be defined as

$$Y := \{(\alpha, \beta) \in F\colon |\alpha| = 1\}, \tag{108.1}$$

and consider the binary relation R on Y, which ist defined by

$$(\alpha, \beta) \, R \, (\gamma, \delta) \;:\Leftrightarrow\; (\alpha = \gamma \wedge \alpha\beta \leq \gamma\delta) \tag{108.2}$$

for all elements (α, β) and (γ, δ) of Y.

To show that the relation is pre-cancellative, let $\mathbf{y} = (\gamma, \delta)$, $\mathbf{y}_1 = (\gamma_1, \delta_1)$ and $\mathbf{y}_2 = (\gamma_2, \delta_2)$ be elements of Y, and $\alpha \in \,]0, 1[$ a scalar such that the vectors

$$\mathbf{y'}_i := \mathbf{y}_i + \alpha (\mathbf{y} - \mathbf{y}_i) := (\gamma'_i, \delta'_i) \tag{108.3}$$

are elements of Y. We have to verify the equivalence

$$\mathbf{y'}_1 \, R \, \mathbf{y'}_2 \;\Leftrightarrow\; \mathbf{y}_1 \, R \, \mathbf{y}_2. \tag{108.4}$$

Now observe that the assumptions $\alpha \in \,]0, 1[$ and $\mathbf{y'}_i \in Y$ imply

$$\gamma = \gamma_i = \gamma'_i \tag{108.5}$$

According to Equation (108.1), we must have

$$|\gamma| = |\gamma_i| = |\gamma'_i| = 1, \tag{108.6}$$

if the vectors $\mathbf{y}$, $\mathbf{y}_i$ and $\mathbf{y'}_i$ are alements of Y. Furthermore, Equation (108.3) implies

$\gamma'_i = \gamma_i + \alpha\,(\gamma - \gamma_i),$ (108.7)

But then the assumption $\gamma_i \neq \gamma$ would lead to $|\gamma'_i| \neq 1$. Since this consideration holds for i = 1 as well as i = 2, we can just write γ for the identical first component of the vectors $\mathbf{y}$, $\mathbf{y}_i$ and $\mathbf{y'}_i$. Then the validity of Equivalence (108.4) results from the following chain:

$$\begin{aligned}
\mathbf{y'}_1\, R\, \mathbf{y'}_2 &\Leftrightarrow (\gamma, \delta'_1)\, R\, (\gamma, \delta'_2) \\
&\Leftrightarrow \gamma\, \delta'_1 \leq \gamma\, \delta'_2 \\
&\Leftrightarrow \gamma\, (\delta_1 + \alpha\, (\delta - \delta_1)) \leq \gamma\, (\delta_2 + \alpha\, (\delta - \delta_2)) \\
&\Leftrightarrow \gamma\, \alpha\, \delta + (1-\alpha)\, \gamma\, \delta_1 \leq \gamma\, \alpha\, \delta + (1-\alpha)\, \gamma\, \delta_2 \\
&\Leftrightarrow \gamma\, \delta_1 \leq \gamma\, \delta_2 \\
&\Leftrightarrow (\gamma, \delta_1)\, R\, (\gamma, \delta_2) \\
&\Leftrightarrow \mathbf{y}_1\, R\, \mathbf{y}_2
\end{aligned}$$

For the transition from the fourth to the fifth line, note that subtracting $\gamma\, \alpha\, \delta$ on both sides of the inequality in the fourth line doesn't change the validity of this inequality, and due to the assumption $\alpha \in\,]0, 1[$, we may also divide by $(1-\alpha)$. So the relation R is pre-cancellative, indeed.

To verify that the same relation is not direction-based, it suffices to give an example, where Equivalence (2.2) is violated, although Equation (2.1) holds with $\lambda > 0$. So consider the vectors

$\mathbf{y}_{11} := (+1, 3),$

$\mathbf{y}_{12} := (+1, 4),$

$\mathbf{y}_{21} := (-1, 5),$

and

$\mathbf{y}_{22} := (-1, 7).$

Obviously, Equation (2.1) holds for $\lambda = 2$. But whereas Equivalence (108.2) leads to $\mathbf{y}_{11}\, R\, \mathbf{y}_{12}$, since $1 \cdot 3 \leq 1 \cdot 4$, the inequality $(-1) \cdot 5 > (-1) \cdot 7$ implies that the relation $\mathbf{y}_{21}\, R\, \mathbf{y}_{22}$ fails to hold. Taken together, these resultsform a violation of Equivalence (2.2).

Exercise 9, p. 93

For given elements $\mathbf{y}_{ij}$, $\mathbf{x}_i$ and $\mathbf{y'}_j$ of the set Y defined by Equation (108.1), let real numbers $\alpha_i \in\,]0, 1[$ be given such that the equation

$\mathbf{y'}_j = \mathbf{y}_{ij} + \alpha_i\, (\mathbf{x}_i - \mathbf{y}_{ij})$ (109.1)

holds for $i \in \{1, 2\}$ and $j \in \{1, 2\}$. Then the arguments leading to Equation (108.5) can be repeated to show that the first components of all vectors $\mathbf{y}_{ij}$, $\mathbf{x}_i$ abd $\mathbf{y'}_j$ must be identical. But this implies that in the example for a violation of direction-basedness in the above solution for Exercise 8 there cannot be any elements $\mathbf{x}_i$ and $\mathbf{y'}_j$ of *Y* such that Equation (109.1) holds for suitable scalars α_i.

More generally, it can be shown for the relation *R* given by Equivalence (108.2) that violations of direction-basedness occur only in situations, where the first components of the vectors $\mathbf{y}_{11}$ and $\mathbf{y}_{12}$ in Equivalence (2.2) are different.

Exercise 10, p. 93

The two claims in the last four lines refer to a subset *Y* of a real vector space with the following property ('solvability condition'): For all elements $\mathbf{y}_{ij}$ of *Y* fullfilling Equation (2.1) for a suitable scalar $\lambda > 0$, there are elements $\mathbf{x}_i$ and $\mathbf{y'}_j$ of *Y* and scalars $\alpha_i \in \,]0, 1[$ such that Equation (109.1) holds for $i \in \{1, 2\}$ and $j \in \{1, 2\}$.

We will first show that the solvability condition holds for every convex subset *Y* of a real vector space *F*. Under this assumption, let $\mathbf{y}_{11}$, $\mathbf{y}_{12}$, $\mathbf{y}_{21}$ and $\mathbf{y}_{22}$ be elements of *Y*, and $\lambda > 0$ a scalar such that Equation (2.1) holds. Furhermore, let γ be an arbitrary element of the interval]0, 1[, and let the scalars α_i and the vectors $\mathbf{x}_i$ be given by Equations (107.1) and (107.3). Then the proof of Equation (107.4) can be taken from Exercise 7. So let $\mathbf{y'}_j$ be the vectors, whose equality is shown in that equation. Then the following chain of implications shows that the introduced scalars and vectors have the properties required by the solvability condition:

$$\lambda > 0 \;\Rightarrow \alpha_i \in \,]0, 1[\;\Rightarrow\; \mathbf{x}_i \in \mathrm{Y} \;\Rightarrow\; \mathbf{y'}_j \in \mathrm{Y}. \tag{110.1}$$

The last two implications in this chain follow from the assumed convexity of *Y*.

Now let *R* be a binary relation on *Y* fulfilling the solvability condition, and we will verify that *R* is direction-based iff it is pre-cancellative. So assume first that *R* is pre-cancellative. and let vectors $\mathbf{y}_{ij}$ and a scalar $\lambda > 0$ be given such that Equation (2.1) holds. To show that Equivalence (2.2) follows from these assumptions, let elements $\mathbf{x}_i$ and $\mathbf{y'}_j$ and scalars α_i be given according to the solvability condition. Then the assumed pre-cancellativeness of the relation *R* implies the equivalence

$$\mathbf{y'}_1 \, R \, \mathbf{y'}_2 \Leftrightarrow \mathbf{y}_{i1} \, R \, \mathbf{y}_{i2} \tag{110.2}$$

for $i \in \{1, 2\}$. Combining these two equivalences, we get Equivalence (2.2).

Conversely, if *R* is direction based, let $\mathbf{y}$, $\mathbf{y}_1$ and $\mathbf{y}_2$ be elements of *Y*, and $\alpha \in \,]0, 1[$ a scalar such that the vectors $\mathbf{y}_{2j}$ given by the definition

$$\mathbf{y}_{2j} := \mathbf{y}_j + \alpha_i\,(\mathbf{y}-\mathbf{y}_j) \tag{110.3}$$

are elements of *Y*. Then Equation (2.1) holds after the additional definitions $\mathbf{y}_{1j} := \mathbf{y}_j$ and $\lambda := 1-\alpha$, and Equivalence (2.2) follows from the assumption that *R* is direction-based. But in the assumed situation, Equivalence (2.3) only reformulates (2.3).

For the claim in the last two lines on p. 93, let *Y* be identical with a real vector space *F*. So *Y* is convex, and we can combine our hitherto obtained results with the final remark in the solution to Exercise 5: A reflexive and transitive relation on *Y* is a linear order iff it is pre-cancencellative.

Exercise 11, p. 95

Let *R* be a direction-based relation on a real vector space *F* with direction cone C_R given by Equation (105.2). We have to show that the relation *R* is pre-Archimedian iff C_R is lineally closed.

So assume first that *R* is pre-Archimedian, and let elements $\mathbf{x}_1$ and $\mathbf{x}_2$ of *F* be given such that $\mathbf{x}_1 \notin C_R$ and $\mathbf{x}_2 \in C_R$. To verify the existence of a scalar δ with the properties specified in Definition A.2.d, observe that Equation (105.2) aloows two rewrite the assumptions $\mathbf{x}_1 \notin C_R$ and $\mathbf{x}_2 \in C_R$ as $\mathbf{0}\, R_c\, \mathbf{x}_1$ and $\mathbf{0}\, R\, \mathbf{x}_2$. Now the assumed pre-Archimedian property of *R* allows to derive the existence of a scalar $\delta > 0$ with the properties given by Definition A.5.c (the roles of the vectors $\mathbf{y}$, $\mathbf{y}_1$ and $\mathbf{y}_2$ in that definition being taken by $\mathbf{0}$, $\mathbf{x}_1$ and $\mathbf{x}_2$ in the present situation). It suffices to show that this δ fulfills the requirements of Definition A.2.d. So let a scalar $\alpha \in\,]0, \delta[$ be given, and note that the definition of the vector $\mathbf{z}_\alpha$ in Equation (A.4) has to be rewritten as

$$\mathbf{z}_\alpha := \mathbf{x}_1 + \alpha\,(\mathbf{x}_2-\mathbf{x}_1). \tag{111.1}$$

for the present situation. Furthermore, since *R* is a relation on the entire vector space *F*, this space takes the role of the set *Y*; hence $\mathbf{z}_\alpha \in Y$. But then the implication $\mathbf{z}_\alpha \in Y \Rightarrow \mathbf{y}\, R_c\, \mathbf{z}_\alpha$ in Definition A.5.c leads to $\mathbf{0}\, R_c\, \mathbf{z}_\alpha$, and another reference to Equation (105.2) gives $\mathbf{x}_1 + \alpha\,(\mathbf{x}_2-\mathbf{x}_1) \notin C_R$.

Conversely, under the assumption that C_R is lineally closed, let $\mathbf{y}$, $\mathbf{y}_1$ and $\mathbf{y}_2$ be elements of *F* such that $\mathbf{y}\, R_c\, \mathbf{y}_1$ and $\mathbf{y}\, R\, \mathbf{y}_2$, and define vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ by $\mathbf{x}_i := \mathbf{y}_i - \mathbf{y}$. Recalling from Corollary 2.2 that a direction-based relation on an entire vector space is invariant under translation, we may apply a translation by $-\mathbf{y}$ to the assumptions $\mathbf{y}\, R_c\, \mathbf{y}_1$ and $\mathbf{y}\, R\, \mathbf{y}_2$, and rewrite them as $\mathbf{0}\, R_c\, \mathbf{x}_1$ resp. $\mathbf{0}\, R\, \mathbf{x}_2$. Using Equation (105.2), we get $\mathbf{x}_1 \notin C_R$, and $\mathbf{x}_2 \in C_R$. Now let δ be a scalar with the property specified in Definition A.2.d, and we will show that this scalar also fulfills the requirements of Definition A.5.c; i.e., for every $\alpha \in]0, \delta[$, the relation $\mathbf{y}\ R\ \mathbf{y}_1 + \alpha\ (\mathbf{y}_2 - \mathbf{y}_1)$ fails to be present. Indeed, this follows from the subsequent chain of equivalences:

$$\mathbf{y}\ R\ \mathbf{y}_1 + \alpha\ (\mathbf{y}_2 - \mathbf{y}_1) \ \Leftrightarrow\ \mathbf{0}\ R\ \mathbf{x}_1 + \alpha\ (\mathbf{x}_2 - \mathbf{x}_1) \ \Leftrightarrow\ \mathbf{x}_1 + \alpha\ (\mathbf{x}_2 - \mathbf{x}_1) \in C_R. \qquad (111.2)$$

For the first equivalence in this chain, we use the invariance under translation (by $-\mathbf{y}$), and the second one is granted by Equation (105.2). But δ has been chosen such that the last link of the chain is false for $\alpha \in]0, \delta[$; so the same holds for the first one.

Exercise 12, p. 97

We will first present a rather informal solution, and then put it into the more formal framework of a probability space. So let $C := \{a, b\}$ be the set of available conditions, X a C-valued random variable indicating the applied experimental condition, and Y a W-valued random variable - the dependent variable of the randomised experiment. Then the following equation holds for every $A \in \mathsf{S}$ in the situation described in the text (P(..) standing for 'probability of ..):

$$\begin{aligned} \mathbf{y}_{ic}(A) &= P(Y \in A) \\ &= P((\mathrm{X} = a \wedge Y \in A) \ \vee\ (X = b \wedge Y \in A)) \\ &= P(X = a) \cdot P(Y \in A \mid X = a) \ +\ P(X = b) \cdot P(Y \in A \mid X = b) \\ &= \alpha \cdot \mathbf{y}_{ia}(A) + (1-\alpha) \cdot \mathrm{y}_{ib}(A). \end{aligned} \qquad (112.1)$$

For a more formal approach, let W be a non-empty set, S a system of subsets of W, and define $C := \{a, b\}$ and $\Omega := C \times W$. Furhtermore, let A_W be the σ-algebra in W generated by the set system S, and let the set C be endowed with its power set $\mathsf{P}W$ as its σ-algebra. Finally, let A be the coursest σ-algebra in Ω, where the projection maps $\mathrm{pr}_1{:}\Omega \to C$ and $\mathrm{pr}_2{:}\Omega \to W$ are measurable, and let P be a probability measure on A. Then the probability space (Ω, A, P) can be considered to model the random selection of an experimental condition and the performance of the experiment under this condition in the following understanding: An element $\omega = (c^*, \mathrm{w})$ of Ω with $c^* \in C$ and $w \in W$ represents the choice of condition c^* and the result w in the experiment. Now let $X{:}\Omega \to C$ and $Y{:}\Omega \to W$ be the two projection maps; i.e., X is a C-valued random variable indicating the selected condition, and Y a W-valued random variable representing the result of the experiment; i.e., X and Y are the independet and the dependent variables of the experiment. In particular, the definition

$\alpha := P(X = a)$ (112.2)

specifies the probability that the experiment is performed under condition a. Without loss of generality, we may assume $0 < \alpha < 1$. (Otherwise, the result is trivial.) Now define the maps $\mathbf{y}_{\mathrm{i}a}:\mathsf{S}\to\mathbb{R}$, $\mathbf{y}_{\mathrm{i}b}:\mathsf{S}\to\mathbb{R}$ and $\mathbf{y}_{\mathrm{i}c}:\mathsf{S}\to\mathbb{R}$ such that the equations

$\mathbf{y}_{\mathrm{i}a}(A) = P(Y \in A \mid X = a)$, (112.3)

$\mathbf{y}_{\mathrm{i}b}(A) = P(Y \in A \mid X = b)$, (112.4)

and

$\mathbf{y}_{\mathrm{i}a}(A) = P(Y \in A)$. (112.5)

Before we proceed, observe that these definitions reflect the verbal statements that the maps $\mathbf{y}_{\mathrm{i}a}$, $\mathbf{y}_{\mathrm{i}b}$ and $\mathbf{y}_{\mathrm{i}c}$ represent the behavioural dispositions of subject i under condition a resp. b resp. in the randomised experiment. Indeed, these statements imply the tacid assumption that the behavioural dispositions of subject i under a given experimental condition a or b are not affected by the randomisation probability α. For this reason, the equation

$\mathbf{y}_{\mathrm{i}c}(A) = \alpha\, \mathbf{y}_{\mathrm{i}a}(A) + (1-\alpha)\, \mathbf{y}_{\mathrm{i}b}(A)$, (112.6)

is introduced only as a plausible assumption and not as a necessary consequence of the assumed situation. To be quite exact, we would have to consider a separate probability measure P_α on A for every $\alpha \in\,]0, 1[$ and introduce the assumption that the equalities

$P_{\alpha'}(Y \in A \mid X = a) = P_{\alpha''}(Y \in A \mid X = a)$ (112.7)

and

$P_{\alpha'}(Y \in A \mid X = b) = P_{\alpha''}(Y \in A \mid X = b)$ (112.8)

hold for all randomisation probabilities α' and α'' and every $A \in \mathsf{S}$. But then we may also maintain the denotation P and let it refer to the probability measure P_α for an arbitrary given $\alpha \in\,]0, 1[$.

In this understanding, Equation (112.1) follows from the above definitions.

Exercise 13, p. 99

Referring to a suggestion in the preliminary remarks, we will first present a sulution under the assumption that S is the power set of a finite set W, and then we will outline a generalisation to other situations. Under the above assumption, let Y be a set of probability measures on S, and $v:W\to\mathbb{R}$ an arbitrary map. Than the definition of the map $g_v:Y\to\mathbb{R}$ in the text can be rewritten as

$g_v(\mathbf{y}) := \sum_{w \in W} v(w) \cdot \mathbf{y}(\{w\})$, (113.1) the validity of this map being required for every $\mathbf{y} \in Y$. But this equation can be easily generalised to the vector space F of all maps $\mathsf{S}\to\mathbb{R}$: We can just require that Equation (113.1) holds for every map $\mathbf{y}:\mathsf{S}\to\mathbb{R}$, and we obtain a map $g_v:F\to\mathbb{R}$, whose linearity is easily verified:

$$\begin{aligned} g_v(\lambda\,\mathbf{y}) &= \textstyle\sum_{w\in W} v(w)\cdot\lambda\cdot\mathbf{y}(\{w\}) \\ &= \lambda\cdot\textstyle\sum_{w\in W} v(w)\cdot\mathbf{y}(\{w\}) \\ &= \lambda\, g_v(\mathbf{y}). \end{aligned} \tag{113.2}$$

Note that the first equality is based on pointwise multiplication of the map $\mathbf{y}$ by the scalar λ. Similarly, pointwise addition of maps $\mathbf{y}_1$ and $\mathbf{y}_2$ leads to the following equation:

$$\begin{aligned} g_v(\mathbf{y}_1+\mathbf{y}_2) &= \textstyle\sum_{w\in W} v(w)\cdot(\mathbf{y}_1(\{w\})+\mathbf{y}_2(\{w\}) \\ &= \textstyle\sum_{w\in W} v(w)\cdot\mathbf{y}_1(\{w\}) + \sum_{w\in W} v(w)\cdot\mathbf{y}_2(\{w\}) \\ &= g_v(\mathbf{y}_1)+g_v(\mathbf{y}_2). \end{aligned} \tag{113.3}$$

So the map $g_v;F\to\mathbb{R}$ isa linear, indeed.

The generalisation of this proof to situations with an infinite set W requires some analytical tools, which may be unfamiliar to some readers. For this proof, let W be a nonempty set, S a system of subsets of W, A the σ-algebra in W which is generated by W, and $v:W\to\mathbb{R}$ a map, which is measurable with respect to A.

Before the proper proof, we should make explicit an assumption underlying the definition of the expectation $\mathsf{E}_{Z\sim\mathbf{y}}\, v(Z)$ on p. 98: The assumption that an element $\mathbf{y}$ of the set Y (i.e., a map $\mathbf{y}:\mathsf{S}\to\mathbb{R}$) 'specifies' the distribution of the random variable Z. More explicitly, this means that every map $\mathbf{y}:\mathsf{S}\to\mathbb{R}$, which is an element of Y, has a unique extension to a probability measure on A. (Otherwise, $\mathbf{y}$ could not specify the distribution of Z.) On the background of this explication, we can introduce the notation $P_{\mathbf{y}}$ for this unique extension.

Now let M_+ be the set of all finite, non-negative measures μ on A, where the integral $\int v\, d\mu$ is finite, and define a map $h:M_+\to\mathbb{R}$ such that $h(\mu)$ is the value of that integral for every $\mu \in M_+$. It was required in the main text that the expectation of the random variable $v(Z)$ must be finite, if Z is a W-valued random variable, whose distribution is specified by an element of Y. In our present notation, we can reformulate this assumption as $P_{\mathbf{y}} \in M_+$. In summary, Y is a set of maps $\mathsf{S}\to\mathbb{R}$, which are restrictions to S of elements of M_+.

On this background, the definition of the map $g_v:Y\to\mathbb{R}$ in the main text can be rewritten as

$$g_v(\mathbf{y}) := h(P_{\mathbf{y}}). \tag{113.4}$$

Having thus reformulated the assumed situation, we will now extend the map $h{:}M_+\to\mathbb{R}$ to a map $h'{:}M\to\mathbb{R}$, where M is the set of those maps $\mu{:}\mathsf{A}\to\mathbb{R}$ which can be represented as a difference of two elements of M. Now the theory of Jordan-Hahn-decompositions (see e.g. Bauer, 1992, p. 125, Exercise 3) allows to specify a unique decomposition of this kind. For every $\mu \in M$, there are unique elements μ_+ and μ_- of M_+ such that the following properties hold for every $A \in \mathsf{A}$:

- $\mu(A) = \mu_+(A) - \mu_-(A)$.
- $\mu_+(A) > 0 \;\Rightarrow \mu_-(A) = 0$.
- $\mu_-(A) > 0 \;\Rightarrow \mu_+(A) = 0$.

The first property can also be formulated as $\mu = \mu_+ - \mu_-$, and this is frequently called the Jordan-Hahn-decomposition of μ. Hence, the map $h'{:}M\to\mathbb{R}$ is well defined by the specification

$$h'(\mu) := h(\mu_+) - h(\mu_-), \qquad (113.5)$$

which must hold for every $\mu \in M$ with the above Jordan-Hahn-decomposition.

An explicit presentation of some further steps of the proof would lead too far; but some readers may verify the following assertions:

- For $\mu \in M_+$, we have $\mu_+ = \mu$; hence h' is an extension of h, indeed.
- The set M is a vector space, or - more precisely - a linear subspace of the space of all maps $\mathsf{A}\to\mathbb{R}$.
- The map h' is linear.
- Let F_0 be the set of those maps $\mathsf{S}\to\mathbb{R}$, which are restrictions to S of elements of M, and define a map $f{:}M\to F_0$ such that $f(\mu)$ (with $\mu \in M$) is the restriction to S of μ. Then F_0 is a linear subspace of F (the set of all maps $\mathsf{S}\to\mathbb{R}$), and the map f is linear and bijective. (The concept of a linear map is explicated in Definition A.2.g only for maps $F\to\mathbb{R}$; but this definition can be generalised to maps from one real vector space into another one.)

On the background of these facts, we may define a map $g{:}F_0\to\mathbb{R}$ by

$$g(\mathbf{y}) := h'(f^{-1}(\mathbf{y})) \qquad (113.6)$$

for every $\mathbf{y} \in F_0$. Again, it is left to the reader whether he or she wants to verify two well known properties og linear maps: The inverse of a bijective linear map is linear, and a concatenation of two linear maps is linear. Hence the map g is linear. At this point, we must introduce a fact, whose proof may overcharge most readers: If F_0 is a linear subspace of a real vector space F, then every linear map $F_0\to\mathbb{R}$ can be extended to a linear map $F\to\mathbb{R}$. (See e.g. Holmes, 1975, p. 3, for a proof.) So let the map $g'{:}F\to\mathbb{R}$ be an arbitrary linear extension of the previously introduced map $g{:}F_0\to\mathbb{R}$. Then it is left to prove that g' is an extension of the map $g_v{:}Y\to\mathbb{R}$ given by Equation (113.4). Now observe that the definition of the map f implies $f^{-1}(\mathbf{y}) = P_{\mathbf{y}}$ for every element $\mathbf{y}$ of Y. Furthermore, since the map h' is an extension of h, the definition of the map g implies $g(\mathbf{y}) = g_v(\mathbf{y})$. Finally, we get $g'(\mathbf{y}) = g_v(\mathbf{y})$, since g' is an extension of g. But if this holds for every $\mathbf{y} \in Y$, then g' is an extension of g_v, indeed.

Exercise 14, p. 99

Let W be the set of possible response patterns in a test consisting of m items with item parameters (real numbers) $\eta_1..\eta_m$. More presisely, every element w of W is an m-dimensional vector with $w_j := 1$, if item j is 'solved', and $w_j = 0$ otherwise. Furthermore, let Y be the set of all probability measures on $\mathsf{P}W$ with the following property: There is a real number θ such that the probability of response pattern w is

$$\prod_{j=1}^{m} \frac{e^{(\theta - \eta_j)\cdot w_j}}{1 + e^{\theta - \eta_j}}$$

It is easily verified that the set Y is convex for m = 1, and we will show that it is non-convex for m > 1. It suffices to present an example of two elements $\mathbf{y}_1$ and $\mathbf{y}_2$ of Y and a scalar $\alpha \in]0, 1[$ such that the vector $\alpha\ \mathbf{y}_1 + (1-\alpha)\ \mathbf{y}_2$ is not contained in Y. For convenience, we introduce a reparametrisation. With the definition $s_j := e^{-\eta_j}$, the set Y is the set of all probability measures on $\mathsf{P}W$, where the probability of response parren w is

$$\prod_{j=1}^{m} \frac{\left(t\cdot s_j\right)^{w_j}}{1 + t\cdot s_j}$$

with $t > 0$ (coressponding to e^{θ} in the former parametrisation.

Now let sets A_j (with j = 1..m) consist of all elements w of W with $w_j = 0$; i.e., the set A_j represents the event 'item j not solved'. Then the equations

$$\mathbf{y}(A_j) = 1 / (1+t\cdot s_j), \tag{114.1}$$

and

$$\mathbf{y}(A_1 \cap A_2) = \mathbf{y}(A_1) \cdot \mathbf{y}(A_2) \tag{114.2}$$

follow immediately for the probability measure $\mathbf{y}$ based on parameter t.

On the background of these equations, it is easy to give an example proving the non-convexity of the set Y. More precisely, we will show that the assumption of a convex set Y would lead to a contradiction. So assume tentatively that Y is convex, let $\mathbf{y}_1$ and $\mathbf{y}_2$ be the probability measures based on parameters $t_1 := 1/s_1$ and $t_2 := 4/s_1$, and with $\alpha := 2/3$, let the vector $\mathbf{y}_3$ be given as

$$\mathbf{y}_3 := \alpha\, \mathbf{y}_1 + (1-\alpha)\, \mathbf{y}_2. \qquad (114.3)$$

For $\mathbf{y}_1$ and $\mathbf{y}_2$, an application of Equation (114.1) with j = 1 leads to $\mathbf{y}_1(A_1) = 0.5$ and $\mathbf{y}_2(A_1) = 0.2$, and then Equation (114.3) yields

$$\mathbf{y}_3(A_1) = \alpha \cdot 0.5 + (1-\alpha) \cdot 0.2 = 0.4. \qquad (114.4)$$

Now the tentatively assumed convexity of Y implies that $\mathbf{y}_3$ is an element of Y. Combining Equations (114.1) and (114.4), we can also reconstruct the underlying parameter t_3 by solving the equation $0.4 = 1 / (1 + t_3 \cdot s_1)$, which leads to $t_3 = 1.5/s_1$.

For an application of Equation (114.1) with j = 2, it is convenient to define $z := s_2/s_1$., which implies $z > 0$, since $s_j > 0$ follows for j = 1..m from the definition of s_j. Then Equation (114.1) leads to

$$\mathbf{y}_1(A_2) = 1 / (1 + t_1 \cdot s_2) = 1 / (1 + z),$$
$$\mathbf{y}_2(A_2) = 1 / (1 + t_2 \cdot s_2) = 1 / (1 + 4\, z),$$

and

$$\mathbf{y}_3(A_2) = 1 / (1 + t_3 \cdot s_2) = 1 / (1 + 1.5\, z).$$

But for $\mathbf{y}_3(A_2)$, we can also apply Equation (114.3) and write

$$\mathbf{y}_3(A_2) = \alpha\, \mathbf{y}_1(A_2) + (1-\alpha)\, \mathbf{y}_2(A_2) = 2 / (3 + 3\, z) + 1 / (3 + 12\, z).$$

Combining the right hand sides of the two equations for $\mathbf{y}_3(A_2)$, we get

$$1 / (1 + 1.5\, z) = 2 / (3 + 3\, z) + 1 / (3 + 12\, z).$$

Now it is easily verified that this equation has exactly two formal solutions: $z = 0$ and $z = 1$. (Multiplying both sides of the equation by the product of the three denominators, we obtain a quadratic equation in z, which canot have more than two solutions for z.) But since the formally correct solution $z = 0$ is excluded by the property $z > 0$, we obtain $z = 1$ as the only solution which is compatible with the assumed situation. So a reference to the definition of z leads to $s_2 = s_1$, and this implies $\mathbf{y}_i(A_2) = \mathbf{y}_i(A_1)$ for i = 1..3.

Finally, we can derive the probabilities $\mathbf{y}_i(A_1 \cap A_2)$ from Equation (114.2):

$$\mathbf{y}_1(A_1 \cap A_2) = 0.5 \cdot 0.5 = 0.25,$$
$$\mathbf{y}_2(A_1 \cap A_2) = 0.2 \cdot 0.2 = 0.04,$$

and

$\mathbf{y}_3(A_1 \cap A_2) = 0.4 \cdot 0.4 = 0.16.$
Obviously, this result contradicts Equation (114.3) with $\alpha = 2/3$.

In summary, we have derived a contradiction from the tentative assumption that the set Y is convex. Hence, Y is non-convex.

Exercise 15, p. 99

It suffices to show that the expectation of the number of 'correct solutions' is a strictly increasing function of the parameter θ in the original parametrisation of the solution to Exercise 14. So let random variables Z_j for j = 1..m be definied such that $Z_j := 1$, if item j is 'solved', and $Z_j := 0$ otherwise. (With the interpretation of w contained in the solution to Exercise 14, we can also say that Z_j is the j^th component of the random variable Z from the main text.) Then it is easily verified that the expectation of Z_j is the probability that item j is solved, and this probability is a strictly increasing function of θ. But obviously, the number of correct solutions is the sum of the Z_j; so the same holds for the respective expections. Finally, the sum of a finite collection of strictly increasing maps is strictly increasing.

Exercise 16, p. 101

Let $\preceq_L$ be the lexicographical order on $\mathbb{R}^2$ as defined in the main text. For a first example, consider the vectors $\mathbf{y} := (5, 3)$, $\mathbf{y}' := (4, 2)$ and $\mathbf{y}'' := (6, 2)$, and the scalar $\delta := 0.5$. Then the property $\mathbf{y}' \prec_L \mathbf{y} \prec_L \mathbf{y}''$ follows immediately from the order of the first components of these vectors. For the vector $\mathbf{z}_\alpha$, which is defined in the main text for $\alpha \in [0, 1]$ by $\mathbf{z}_\alpha := \mathbf{y}' + \alpha\,(\mathbf{y}'' - \mathbf{y}')$, we can write $\mathbf{z}_\alpha = (4 + 2\,\alpha, 2)$ in this situation. Now the first component of this vector is smaller than 5 for $0 \leq \alpha < \delta$; so $\mathbf{z}_\alpha \prec_L \mathbf{y}$ follows from the definition of the relation $\preceq_L$. For $\alpha = \delta$, we have $\mathbf{z}_\alpha = (5, 2)$, and since the first components of this vector and of $\mathbf{y}$ are identical, their order is determined by the second component, leading to $\mathbf{z}_\alpha \prec_L \mathbf{y}$. Finally, for $\delta < \alpha \leq 1$, the first component of $\mathbf{z}_\alpha$ is greater than 5, leading to $\mathbf{y} \prec_L \mathbf{z}\alpha$. In summary, we have $\mathbf{z}_\alpha \prec_L \mathbf{y}$ for $\alpha \in [0, \delta]$, and $\mathbf{y} \prec_L \mathbf{z}_\alpha$ for $\alpha \in \,]\delta, 1]$.

A second example is identical with the first one with the exception $\mathbf{y} := (5, 1)$. For $\alpha \neq \delta$, the arguments of the first example can be repeated. For $\alpha = \delta$, we have again $\mathbf{z}_\alpha = (5, 2)$, and the order of the vectors $\mathbf{z}_\alpha$ and $\mathbf{y}$ is determined by their second components; but for the new vector $\mathbf{y}$ this means $\mathbf{y} \prec_L \mathbf{z}_\alpha$. So the changed definition of $\mathbf{y}$ leads to $\mathbf{z}_\alpha \prec_L \mathbf{y}$ for $\alpha \in [0, \delta[$, and $\mathbf{y} \prec_L \mathbf{z}_\alpha$ for $\alpha \in [\delta, 1]$.

Exercise 17, p. 103

The largest abvantage of $\mathbf{y}$ over $\mathbf{z}$ is in the first dimension; but it is smaller than the advantage of $\mathbf{z}$ over $\mathbf{y}$ in the third dimension, leading to $\mathbf{y} \prec \mathbf{z}$. Finally, the largest advantage of $\mathbf{x}$ over $\mathbf{z}$ (first dimension) is greater than the largest advantage of $\mathbf{z}$ over $\mathbf{x}$ (third dimension); hence $\mathbf{z} \prec \mathbf{x}$

Exercise 18, p. 103

In this exercise and the following one, the default range 1..2 of subscript j is replaced by 1..3.

To generalise the order based on the 'largest advantage' of one vector over another one, we should start with a more formal explication of the notion of largest advantages. For an ordered pair $(\mathbf{x}_1, \mathbf{x}_2)$ of elements of $\mathbb{R}^3$ with components $\mathbf{x}_1(j)$ resp. $\mathbf{x}_2(j)$, we define real numbers $d_j(\mathbf{x}_1, \mathbf{x}_2)$ as follows: Let $\{k_1, k_2, k_3)$ be a permutation of the numbers 1, 2 and 3 such that

$$\mathbf{x}_2(k_1) - \mathbf{x}_1(k_1) \le \mathbf{x}_2(k_2) - \mathbf{x}_1(k_2) \le \mathbf{x}_2(k_3) - \mathbf{x}_1(k_3). \quad (118.1)$$

Then the three differences in this inequality are called $d_j(\mathbf{x}_1, \mathbf{x}_2)$; i.e.:

$$d_j(\mathbf{x}_1, \mathbf{x}_2) := \mathbf{x}_2(k_j) - \mathbf{x}_1(k_j). \quad (118.2)$$

Note that the permutation $\{k_1, k_2, k_3)$ will not be unique in situatios, where some of the differences in inequality (118.1) are equal; but even in such situations, the numbers $d_j(\mathbf{x}_1, \mathbf{x}_2)$ given by Equation (118.2) are independet of the choice of a particular permutation from those fulfilling Inequality (118.1).

In this notation, the largest advantage of $\mathbf{x}_1$ over $\mathbf{x}_2$ is $-d_1(\mathbf{x}_1, \mathbf{x}_2)$, whereas the largest advantage of $\mathbf{x}_2$ over $\mathbf{x}_1$ is $d_3(\mathbf{x}_1,\mathbf{x}_2)$.

In the main text, two preference relations are defined, which differ in the treatment of situations with $-d_1(\mathbf{x}_1, \mathbf{x}_2) = d_3(\mathbf{x}_1,\mathbf{x}_2)$. These preference relations will subsequently be denoted as $\preceq'$ and $\preceq''$. The following equivalences have to be understood as definitions applying to all vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ of $\mathbb{R}^3$:

$\mathbf{x}_1 \preceq' \mathbf{x}_2 \;:\Leftrightarrow\; -d_1(\mathbf{x}_1, \mathbf{x}_2) \leq d_3(\mathbf{x}_1, \mathbf{x}_2).$ (118.3)

$\mathbf{x}_1 \preceq'' \mathbf{x}_2 \;:\Leftrightarrow\; -d_1(\mathbf{x}_1, \mathbf{x}_2) < d_3(\mathbf{x}_1, \mathbf{x}_2) \;\vee$

$(-d_1(\mathbf{x}_1, \mathbf{x}_2) = d_3(\mathbf{x}_1, \mathbf{x}_2) \wedge d_2(\mathbf{x}_1, \mathbf{x}_2 \geq 0).$ (118.4)

The reader is invited to reformulate in this notation the arguments leading to $\mathbf{x} \prec \mathbf{y}$, $\mathbf{y} \prec \mathbf{z}$, and $\mathbf{z} \prec \mathbf{x}$ in themain text and in the solution to Exercise 17. So the relations $\preceq'$ and $\preceq''$ aren't orders, since they are intransitive. Nevertheless, notations like $\mathbf{x} \prec' \mathbf{y}$ have to be understood as $\mathbf{x} \preceq' \mathbf{y} \wedge \neg(\mathbf{y} \preceq' \mathbf{x})$. Similarly, $\mathbf{x} \prec'' \mathbf{y}$ means $\mathbf{x} \preceq'' \mathbf{y} \wedge \neg(\mathbf{y} \preceq'' \mathbf{x})$.

Having thus defined the relations, we can easily show that they are direction-based. For vectors $\mathbf{y}_{11}$, $\mathbf{y}_{12}$, $\mathbf{y}_{21}$ and $\mathbf{y}_{22}$ and a scalar $\lambda > 0$ with the property $\mathbf{y}_{22}-\mathbf{y}_{21} = \lambda\,(\mathbf{y}_{12}-\mathbf{y}_{11})$, the equation $d_j(\mathbf{y}_{21}, \mathbf{y}_{22}) = \lambda\, d_j(\mathbf{y}_{11}, \mathbf{y}_{12})$ follows immediately (for j = 1...3) from the definition of d_j in Equation (118.2). But then the definitions of the relations $\preceq'$ and $\preceq''$ immediately lead to the equivalences $\mathbf{y}_{11} \preceq' \mathbf{y}_{12} \Leftrightarrow \mathbf{y}_{21} \preceq' \mathbf{y}_{22}$ and $\mathbf{y}_{11} \preceq'' \mathbf{y}_{12} \Leftrightarrow \mathbf{y}_{21} \preceq'' \mathbf{y}_{22}$.

Exercise 19, p. 104

The vectors **x**, **y** and **z** introduced in the main text show that the relations $\preceq'$ and $\preceq''$ are intransitive. Hence, the non-convexity of the respective direction cones follows from Corollary 3.1.(5).

Exercise 20, p. 104

Let **x**, **y** and **z** be element of a real vectorspace *F*. Now consider the equation

$\mathbf{a} + \alpha\,(\mathbf{c}-\mathbf{a}) = \mathbf{x} + \lambda\,(\mathbf{b}-\mathbf{x})$ (120.1)

immediately before Footnote 16 of the main text , which is the main part of an explication of the relation 'viewed from **x**, vector **b** lies between **a** and **c**'. For the vectors **x**, **y** and **z** of the present situation, we have to write the relation 'viewed from **0**, the vector **z**−**x** lies between **y**−**x** and **z**−**y**'. So we have to assign the roles of the vectors **a**, **b**, **c** and **x** in the above equation to the vectors **y**−**x**, **z**−**x**, **z**−**y** and **0**, leading to

$(\mathbf{y}-\mathbf{x}) + \alpha\,((\mathbf{z}-\mathbf{y}) - (\mathbf{y}-\mathbf{x})) \;=\; \mathbf{0} + \lambda\,((\mathbf{z}-\mathbf{x}) - \mathbf{0}).$ (120.2)

Taking $\alpha = \lambda = 0.5$, and using well known rules for bracket-operations, we get

$0.5\,\mathbf{z} - 0.5\,\mathbf{y} \;= 0.5\,\mathbf{z} - 0.5\,\mathbf{y},$ (120.3) and no furhter proof is necessary to verify the validity of this equation for all elements **x**, **y** and **z** of a real vector space *F*.

Nevertheless, some readers may ask whether all algebraic operations known from real numbers are legal in real vector spaces. Indeed, the answer to this question is negative. For readers of limited familiarity with real vector spaces, it may be a good exercise to justify systematically all operations, which have been applied above, by the axioms of real vector spaces. (See Definition A.1 in the appendix of the main text for these axioms.)

Exercise 21, p. 106

In the context of p. 106, the set C_R is defined as the smallest cone in F containing D_R as a subset. To verify the equation

$$C_R = \cup_{\lambda>0} \lambda D_R, \tag{121.1}$$

it suffices to prove the following general proposition, which will also be helpful in further Exercises:

Propostion 121.1.1 Let A be a non-empty subset of a real vector space F, and let a subset C of F be defined as

$$C := \cup_{\lambda>0} \lambda A. \tag{121.2}$$

Then C is the smallest cone in F containing A as a subset.

For a solution of Exercise 21, we identify the set D_R from the main text with the set A of the above proposition. Then the proposition will imply that C_R (being defined as the smallest cone in F containing D_R as a subset) is identical with the set C given by Equation (121.2).

Of course, the product λA of a scalar λ and a subset A of a real vector space has to be interpreted in the undserstanding of the respective notational convention (immediatly after Definition A.1 in the appendix of the main text): It is the set of all vectors λ **a**, where **a** is an element of A. In particular for $\lambda = 1$, we obtain $\lambda A = A$. So A is a subset of C, and it is left to show that C is a cone, and a subset of every cone containing A as a subset.

To prove these properties, let $C^{\sim}$ be an arbitrary cone in F containing A as a subset, **x** an arbitrary element of C, and λ a strictly positive scalar. It suffices to verify the properties $\mathbf{x} \in C^{\sim}$ and $\lambda\, \mathbf{x} \in C$. So let **a** be an element of A, and μ a strictly positive scalar such that $\mathbf{x} = \mu\, \mathbf{a}$, the existence of **a** and

μ with these properties being granted by the assumption $\mathbf{x} \in C$ and by the definition of C. Then $\mathbf{a}$ is an element of $C^\sim$, since A is a subset of $C^\sim$, and this implies $\mu\, \mathbf{a} \in C^\sim$, since $C^\sim$ is a cone. Finally, the equation $\lambda\, \mathbf{x} = (\lambda{\cdot}\mu)\, \mathbf{a}$ leads to $\lambda\, \mathbf{x} \in C$, since $\lambda \cdot \mu > 0$.

Exercise 22, p. 106

We have to verify the claim that the cone C_R (i.e., the smallest cone in F containing D_R) is a potential direction cone of the relation R. In other words, we have to show that the equivalence

$$\mathbf{y}_1\, R\, \mathbf{y}_2 \;\Leftrightarrow \mathbf{y}_2 - \mathbf{y}_1 \in C_R \tag{122.1}$$

holds for all elements $\mathbf{y}_1$ and $\mathbf{y}_2$ of Y.

So let $\mathbf{y}_1$ and $\mathbf{y}_2$ be arbitrary elements of Y, and assume for the proof of the forward implication that the relation $\mathbf{y}_1\, R\, \mathbf{y}_2$ holds. Then the vector $\mathbf{y}_2 - \mathbf{y}_1$ is an element of D_R by the definition of this set. So this difference is also an element of the cone C_R , which contains D_R as a subset.

Conversely, if $\mathbf{y}_2{-}\mathbf{y}_1 \in C_R$, let $\mathbf{x}$ be an element of D_R and λ a strictly positive scalar such that $\mathbf{y}_2{-}\mathbf{y}_1 = \lambda\, \mathbf{x}$. (The existence of such λ and $\mathbf{x}$ granted by the equation $C_R = \bigcup_{\lambda>0} \lambda\, D_R$, which has been established in Exercise 23.) Then the Definition of D_R implies the existence of elements $\mathbf{y'}_1$ and $\mathbf{y'}_2$ of Y such that $\mathbf{x} = \mathbf{y'}_2{-}\mathbf{y'}_1$ and $\mathbf{y'}_1\, R\, \mathbf{y'}_2$. But for a direction-based relation R, this implies $\mathbf{y}_1\, R\, \mathbf{y}_2$, since the assumed properties can be combined to $\mathbf{y}_2{-}\mathbf{y}_1 = \lambda\, (\mathbf{y'}_2{-}\mathbf{y'}_1)$.

Exercise 23, p. 106

We have to show that C_R is the only potential direction cone of the relation R iff $C_D = F$. A proof with a uselful side result can be based on the equation

$$C^\sim \cap C_D = C_R, \tag{123.1}$$

which has been introduced in the main text for every cone $C^\sim$ in F as a necessary and sufficient condition of being a potential direction cone. With the notation D for the set system of all potential direction cones, this set system is described by the equation

$$\mathsf{D} = \{C^\sim \subseteq F\colon \exists\, A \subseteq (F\backslash C_D)\colon C^\sim = C_R \cup \textstyle\bigcup_{\lambda>0} \lambda\, A\}. \tag{123.2}$$

Proof: For every potential direction cone $C^\sim$, we can take the set $A := C^\sim \backslash C_D$. Conversely, if A is a subset of $F\backslash C_D$, then Proposition 121.1 shows that the set $\bigcup_{\lambda>0} \lambda\, A$ is a cone. It can be left to the reader to show that the union of two cones (e.g. C_R and $\bigcup_{\lambda>0} \lambda A$) is a cone and that Equation (123.1) holds for every element $C^{\sim}$ of the set system D specified by Equation (123.2).

Recalling that C_R is a subset of C_D, we see that every subset A of $F \backslash C_D$ is disjoint from C_R. On this background, Equation (123.2) allows the following conclusion: The set C_R is the only element of the set system D iff the empty set is the only subset A of $F \backslash C_D$. Obviously, this condition is equivalent with $C_D = F$, since C_D is a subset of F.

Exercise 24, p. 106

We have to show that the sets C_R, which are analysed in the context of p. 106 and in Corollary 3.1 are identical under the assumption $Y = F$. To avoid ambiguity of notation, let $C^{\sim}$ be a potential direction cone of the relation R (whose uniqueness in situations with $Y = F$ is shown on p. 106), and reserve the denotation C_R for the cone treated in Corollary 3.1; i.e.,

$$C_R := \{\mathbf{x} \in F\colon \mathbf{0}\ R\ \mathbf{x}\}. \tag{124.1}$$

Now let $\mathbf{x}$ be an arbitrary element of F, and we will verify the equivalence

$$\mathbf{x}-\mathbf{0} \in C^{\sim} \Leftrightarrow \mathbf{x} \in C_R. \tag{124.2}$$

For the forward implication, the assumption $\mathbf{x}\text{-}\mathbf{0} \in C^{\sim}$ immediately implies $\mathbf{0}\ R\ \mathbf{x}$, since $C^{\sim}$ is a potential direction cone. So $\mathbf{x} \in C_R$ is granted by Equation (124.1). Conversely, for $\mathbf{x} \in C_R$, Equation (124.1) gives $\mathbf{0}\ R\ \mathbf{x}$, and $\mathbf{x}-\mathbf{0}$ follows, since $C^{\sim}$ is a potential direction cone.

Exercise 25, p. 107

Let $C^{\sim}$ be a potential direction cone of a direction based relation R on a non-empty subset Y of a real vector space F. We have to show that the cone $C^{\sim}$ is pointed (i.e., $\mathbf{0} \in C^{\sim}$) iff the relation R is reflexive.

Under the assumption $\mathbf{0} \in C^{\sim}$, we have $\mathbf{y}-\mathbf{y} \in C^{\sim}$ for every element $\mathbf{y}$ of the set Y, and this implies $\mathbf{y}\ R\ \mathbf{y}$, since $C^{\sim}$ is assumed to be a potential direction cone. But if $\mathbf{y}\ R\ \mathbf{y}$ holds for every $\mathbf{y} \in Y$, then the relation R is reflexive.

Conversely, the assumption of a reflexive relation R implies $\mathbf{y}\ R\ \mathbf{y}$ for every element $\mathbf{y}$ of Y, which implies $\mathbf{y}-\mathbf{y} \in C^{\sim}$ for the direction cone $C^{\sim}$.

Exercise 26, p. 108

To verify the equivalence of Assertions (1) and (3) of Corollary 3.2, we will first present an easy proof of the implication (3) ⇒ (1). So assume that the direction cone C_R of a direction-based relation R on a convex set Y is convex, and let $\mathbf{y}_0$, $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$ be elements of Y such that the equation

$$\mathbf{y}_1 + \alpha\,(\mathbf{y}_3 - \mathbf{y}_1) = \mathbf{y}_0 + \lambda\,(\mathbf{y}_2 - \mathbf{y}_0) \tag{126.1}$$

holds for suitable real numbers $\alpha \in\,]0, 1[$ and $\lambda > 0$. To verify the implication concluding Assertion (1), we have to derive the relation $\mathbf{y}_0\ R\ \mathbf{y}_2$ from the premissa the relations $\mathbf{y}_0\ R\ \mathbf{y}_1$ and $\mathbf{y}_0\ R\ \mathbf{y}_3$ hold. Now this premissa implies that the vectors $\mathbf{y}_1 - \mathbf{y}_0$ and $\mathbf{y}_3\text{-}\mathbf{y}_0$ are elements of the direction cone C_R. Furthermore, Equation (126.1) can be rewritten as

$$\lambda\,(\mathbf{y}_2 - \mathbf{y}_0) = \alpha\,(\mathbf{y}_3 - \mathbf{y}_0) + (1-\alpha)\,(\mathbf{y}_1 - \mathbf{y}_0). \tag{126.2}$$

Since the right hand side of this equation is a convex linear combination of elements of the convex set C_R, the vector $\lambda\,(\mathbf{y}_2 - \mathbf{y}_0)$ is also an element of this set, and this implies $\mathbf{y}_2 - \mathbf{y}_0 \in C_R$, since C_R is a cone. But then $\mathbf{y}_0\ R\ \mathbf{y}_2$ follows from the property defining a direction cone.

To verify the implication (1)⇒(3), assume that Assertion (1) holds, let $\mathbf{x}_1$ and $\mathbf{x}_2$ be arbitrary elements of the direction cone C_R, α a scalar with $0 < \alpha < 1$. We have to verify $\mathbf{x} \in C_\mathrm{R}$ for the vector $\mathbf{x}$ defined as

$$\mathbf{x} := \alpha\,\mathbf{x}_2 + (1-\alpha)\,\mathbf{x}_1. \tag{126.3}$$

In a first step, we will prove the existence of a scalar $\lambda > 0$ and of elements $\mathbf{y}_{\mathrm{ij}}$ of Y such that $\mathbf{y}_{\mathrm{i}1}\ R\ \mathbf{y}\mathrm{i}2$, and $\mathbf{x}_\mathrm{i} = \lambda\,(\mathbf{y}_{\mathrm{i}2} - \mathbf{y}_{\mathrm{i}1})$. Equation (121.1) implies the existence of strictly positive scalars λ_1 and λ_2 and of elements $\mathbf{x'}_1$ and $\mathbf{x'}_2$ of D_R such that $\mathbf{x}_\mathrm{i} = \lambda_\mathrm{i}\,\mathbf{x'}_\mathrm{i}$. The derivation of a scalar λ and elements $\mathbf{y}_{\mathrm{ij}}$ of Y depends on the order of λ_1 and λ_2.

- For $\lambda_1 \le \lambda_2$, define $\lambda := \lambda_2$, $\beta := \lambda_1/\lambda_2$, and let $\mathbf{y}$, $\mathbf{y}_{12}$, $\mathbf{y}_{21}$ and $\mathbf{y}_{22}$ be elements of Y such that $\mathbf{x'}_1 = \mathbf{y}_{12} - \mathbf{y}$, $\mathbf{x'}_2 = \mathbf{y}_{22} - \mathbf{y}_{21}$, $\mathbf{y}\ R\ \mathbf{y}_{12}$, and $\mathbf{y}_{21}\ R\ \mathbf{y}_{22}$, the existence of elements of Y with these properties being granted by the definition of the set D_R. Then the definition

  $$\mathbf{y}_{11} := \beta\,\mathbf{y} + (1-\beta)\,\mathbf{y}_{12}, \tag{126.4}$$

  gives another element of Y (recall the assumed convexity of this set) such that

  $$\mathbf{y}_{12} - \mathbf{y}_{11} = \beta\,(\mathbf{y}_{12} - \mathbf{y}). \tag{126.5}$$

  This equation has two consequences. First, $\mathbf{y}_{11}\ R\ \mathbf{y}_{12}$ follows from $\mathbf{y}\ R\ \mathbf{y}_{12}$, since R is direction-based, and $\beta > 0$. Furthermore, the assumptions, definitions and results can be combined in the equations

$\mathbf{x}_1 = \lambda_1\, \mathbf{x'}_1 = \lambda_2 \cdot \beta\, (\mathbf{y}_{12} - \mathbf{y}) = \lambda\, (\mathbf{y}_{12} - \mathbf{y}_{11})$, (126.6)

and

$\mathbf{x}_2 = \lambda_2\, \mathbf{x'}_2 = \lambda\, (\mathbf{y}_{22} - \mathbf{y}_{21})$. (126.7)

- For $\lambda_1 > \lambda_2$, we use a symmetric approach. Define $\lambda := \lambda_1$, $\beta := \lambda_2/\lambda_1$, and let $\mathbf{y}_{11}$, $\mathbf{y}_{12}$, $\mathbf{y}$ and $\mathbf{y}_{22}$ be elements of Y with the properties $\mathbf{x'}_1 = \mathbf{y}_{12} - \mathbf{y}_{11}$, $\mathbf{x'}_2 = \mathbf{y}_{22} - \mathbf{y}$, $\mathbf{y}_{11}\, R\, \mathbf{y}_{12}$ and $\mathbf{y}\, R\, \mathbf{y}_{22}$. After the definiton

  $\mathbf{y}_{21} := \beta\, \mathbf{y} + (1-\beta)\, \mathbf{y}_{22}$, (126.8)

  we can repeat the above arguments to verify the properties $\mathbf{x}_\mathrm{i} = \lambda\, (\mathbf{y}_{\mathrm{i}2} - \mathbf{y}_{12})$ and $\mathbf{y}_{\mathrm{i}1}\, R\, \mathbf{y}_{\mathrm{i}2}$.

Having established the existence of a scalar $\lambda > 0$ and elements $\mathbf{y}_{\mathrm{ij}}$ with these properties, we define further vectors as suggested in Footnote 20 of the main text.:

$\mathbf{y}_0 := 0.25\, (\mathbf{y}_{11} + \mathbf{y}_{12} + \mathbf{y}_{21} + \mathbf{y}_{22})$, (126.9)

$\mathbf{y}_1 := \mathbf{y}_0 + 0.25\, (\mathbf{y}_{12} - \mathbf{y}_{11})$, (126.10)

$\mathbf{y}_3 := \mathbf{y}_0 + 0.25\, (\mathbf{y}_{22} - \mathbf{y}_{21})$, (126.11)

and

$\mathbf{y}_2 := \mathbf{y}_1 + \alpha\, (\mathbf{y}_3 - \mathbf{y}_1)$. (126.12)

Certainly, these vectors are elements of Y due to the convexity of this set. Furthermore, the definitions imply

$\mathbf{y}_1 - \mathbf{y}_0 = 0.25\, (\mathbf{y}_{12} - \mathbf{y}_{11})$ (126.13)

and

$\mathbf{y}_3 - \mathbf{y}_0 = 0.25\, (\mathbf{y}_{22} - \mathbf{y}_{21})$, (126.14)

and these equations allow to derive $\mathbf{y}_0\, R\, \mathbf{y}_1$ and $\mathbf{y}_0\, R\, \mathbf{y}_3$ from $\mathbf{y}_{11}\, R\, \mathbf{y}_{12}$ resp. $\mathbf{y}_{21}\, R\, \mathbf{y}_{22}$, since the relation R is direction-based. So the assumed validity of Assertion (1) leads to $\mathbf{y}_0\, R\, \mathbf{y}_2$, and this is equivalent with $\mathbf{y}_2 - \mathbf{y}_0 \in C_R$, since C_R is the direction cone of the relation R. But then $\mathbf{x} \in C_R$ follows from the cone property of C_R and the subsequente contionuation of Equation (126.3):

$$\begin{aligned}
\mathbf{x} &= \alpha\, \mathbf{x}_2 + (1-\alpha)\, \mathbf{x}_1. \\
&= \alpha \cdot \lambda\, (\mathbf{y}_{22} - \mathbf{y}_{21}) + (1-\alpha) \cdot \lambda\, (\mathbf{y}_{12} - \mathbf{y}_{11}) \\
&= \lambda\, (\alpha \cdot 4\, (\mathbf{y}_3 - \mathbf{y}_0) + (1-\alpha) \cdot 4\, (\mathbf{y}_1 - \mathbf{y}_0)) \\
&= 4\lambda\, (\mathbf{y}_1 + \alpha\, (\mathbf{y}_3 - \mathbf{y}_1) - \mathbf{y}_0) \\
&= 4\lambda\, (\mathbf{y}_2 - \mathbf{y}_0). \qquad (126.15)
\end{aligned}$$

Note that the transitions to the third line and to the last one are based on the definitions of the vectors $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_2$.

Exercise 27, p. 109

We may use Corollary 2.2 and transfer the roles of the relation R and the cone C_R of that corollary to the relation $R^{\sim}$ and the cone $C^{\sim}$ of the present situation.

Exercise 28, p. 109

In the situation assumed on p. 109 of the main text, the direction cone C_R of the relation R is also the direction cone of the relation R' (see Exercise 27). For situations, where the smallest affine subspace of F containing Y is a proper subset of F, we have to show that some element of F cannot be represented as a difference of two elements of C_R.

So let R be a direction-based relation on a non-empty subset Y of a real vector space F. Furthermore, let X^* be the smallest affine subspace of F including the set Y as a subset, assume that and let X be a linear subspace of F resulting from a suitable translation of X^*. In other words, X is a linear subspace of F such that the equation

$$X = \{\mathbf{x} + \mathbf{x}_t : \mathbf{x} \in X^*\} \qquad (128.1\}$$

holds for a suitable translation vector $\mathbf{x}_t \in F$. So let $\mathbf{x}_t$ be a given element of F with this property. We claim that the direction cone C_R of the relation R is a subset of X. If this claim can be verified, then all differences of elements of C_R must be elements of X, since X is a linear subspace.[1] Hence, for $\mathbf{x} \in F \backslash X^*$, the vector $\mathbf{x} + \mathbf{x}_t$ cannot be represented as a difference of two elements of C_R, since $\mathbf{x} + \mathbf{x}_t \notin X$ follows from $\mathbf{x} \in F \backslash X^*$ by Equation (128.1).

To prove the inclusion $C_R \subseteq X$, let $\mathbf{x}$ be an arbitrary element of C_R, and we will show that $\mathbf{x}$ is also an element of X. So let $\mathbf{y}_1$ and $\mathbf{y}_2$ be elements of Y, and λ a scalar such that

---

[1] Although Definition A.2.e requires only that a linear subspace is closed under addition and scalar multiplication, this implies being closed under differences: If $\mathbf{x}_1$ and $\mathbf{x}_2$ are elements of a linear subspace S of a real vector space, then the difference $\mathbf{x}_2 - \mathbf{x}_1$ can be written as $\mathbf{x}_2 + \lambda\,\mathbf{x}_1$ with $\lambda = -1$. But $\lambda\,\mathbf{x}_1 \in S$ follows from closure under scalar multiplication, and then closure under addition leads to $\mathbf{x}_2 + \lambda\,\mathbf{x}_1 \in S$.

Readers, who are uncertain whether the property $(-1) \cdot \mathbf{x} = -\mathbf{x}$ (which is well known in $\mathbb{R}^n$) applies to general real vector spaces, can derive it from Assertion (5) of Definition A.2, using $\lambda := -1$ and $\mu = 1$

$\mathbf{x} = \lambda\,(\mathbf{y}_2 - \mathbf{y}_1).$ (128.2)
(For the existence of $\mathbf{y}_1$, $\mathbf{y}_2$ and λ with these properties, note that Equation (121.1) implies the existence of a scalar λ and an element $\mathbf{x}'$ of the set D_R such that $\mathbf{x} = \lambda\,\mathbf{x}'$. Furthermore, the existence of elements $\mathbf{y}_1$ and $\mathbf{y}_2$ of Y with $\mathbf{x}' = \mathbf{y}_2 - \mathbf{y}_1$ follows from the definition of D_R on p. 105.)

Now observe that the vectors $\mathbf{y}_1$ and $\mathbf{y}_2$ - being elements of Y - must also be contained in X^*, since X^* is assumed to include Y as a subset. So the vectors $\mathbf{y}_1 + \mathbf{x}_t$ and $\mathbf{y}_2 + \mathbf{x}_t$ are elements of X by Equation (128.1). But then X - being a linear subspace of F - must also contain the vector $\lambda\,((\mathbf{y}_2 + \mathbf{x}_t) - (\mathbf{y}_1 + \mathbf{x}_t))$, which is equal to $\mathbf{x}$ by Equation (128.2).

In fact, the inclusion $C_R \subseteq X$ could also be taken from Lemma A.6.(3). Note, however, that both sets are defined in Lemma A.6 in a way differing formally from their introduction in the context of p. 109. Although the equivalence of the definitions can be derived from 'well known' properties of linear and affine subspaces, a version of this proof, which is comprehensible for beginners in the area of general real vector spaces, would be rather tedious.

References:

Bauer, H. (1992). *Maß- und Integrationstheorie* (2nd, revised edition). Berlin: De Gruyter.
Holmes, R.B. (1975). *Geometric Functional Analysis and its Applications.* New York: Springer.