Introduction: Exercises

1.1 (a, b) The two formal definitions are as follows.

\((\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

\((\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

where \(\varepsilon, \delta, a, x\) range over \(\mathbb{R}\). The crucial difference is the position of \(a\). For continuity the required \(\delta\) depends on \(a\) and \(\varepsilon\). For uniform continuity the required \(\delta\) depends only on \(\varepsilon\), it must work for all \(a\).

(c) Using the formal definition of continuity we see that

\(\neg(\forall a)(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

will give the required formal definition. We simplify by taking \(\neg\) though the sentence to the atomic parts. Here are the steps for this.

\(\neg(\forall a)(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

\((\exists a)(\neg(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

\((\exists a)(\exists \varepsilon > 0)(\neg(\exists \delta > 0)(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

\((\exists a)(\exists \varepsilon > 0)(\forall \delta > 0)(\neg(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

\((\exists a)(\exists \varepsilon > 0)(\forall \delta > 0)(\exists x)[|x - a| < \delta \text{ and } |f(x) - f(a)| \neq \varepsilon]\)

which we can rephrase as

For some \(a\) there is some sufficiently small interval

\[I = [f(a) - \varepsilon, f(a) + \varepsilon]\]

around \(f(a)\), (for some \(\varepsilon > 0\)) such that each interval around \(a\)

\[(a - \delta, a + \delta)\]

(for \(\delta > 0\)) contains an elements outside \(I\).

which, in a way, is not as understandable as the formal definition.

(d) Using the same trick as in (c) we have

\(\neg(\forall \varepsilon > 0)(\exists \delta > 0)(\forall a)(\forall x)[|x - a| < \delta \implies |f(x) - f(a)| < \varepsilon]\)

which becomes

\((\exists \varepsilon > 0)(\forall \delta > 0)(\exists a)(\exists x)[|x - a| < \delta \text{ and } |f(x) - f(a)| \neq \varepsilon]\)

which needs to be read at least a couple of times. ■
The axiom of choice defined: Exercises

2.1 When there is just one set, \((A_0)\) we take any element \(a \in A_0\) and then the rather trivial function

\[
\begin{array}{ccc}
0 & \rightarrow & A_0 \\
\{0\} & \rightarrow & a
\end{array}
\]

is a choice function. Note that a selection of \(a \in A_0\) does not need a use of choice.

We now proceed by induction on \(n\). Thus for the induction step, \(n \mapsto n + 1\), consider a family \((A_0, \ldots, A_n)\) of non-empty sets together with an extra non-empty set \(A_{n+1}\). We may assume given (by an earlier construction) a function

\[
f : \{0, \ldots, n\} \rightarrow A_0 \cup \cdots \cup A_n
\]

with

\[
f(i) \in A_i
\]

for each index \(0 \leq i \leq n\). Now consider any element \(a \in A_{n+1}\), and let

\[
g : \{0, \ldots, n, n + 1\} \rightarrow A_0 \cup \cdots \cup A_n \cup A_{n+1}
\]

be the function given by

\[
g(i) = \begin{cases} 
a & \text{if } i = n + 1 \\
f(i) & \text{if } 0 \leq i \leq n
\end{cases}
\]

for each index \(0 \leq i \leq n + 1\). This is a choice function for the extended family.

2.2 Let us first show that there is at least one such function \(f\).

For any such function \(f\), for each \(b \in B\) we require \(f(b) \in A\) so that \(f(b)\) must be choice function for the given indexed set. Thus we require

\[
f(b)(i) \in A_i
\]

for each index \(i\). Thus we may set

\[
f(b)(i) = f_i(b)
\]

to obtain a function \(f\) of the correct type (that is \(f\) moves from \(B\) to \(A\)). With this, for each index \(i\) and element \(b \in B\), we have

\[
(p_i \circ f)(b) = p_i(f(b)) = f(b)(i) = f_i(b)
\]

so that

\[
p_i \circ f = f_i
\]

to show that all the required triangles do commute.

To show that this is the only possible function suppose \(g\) is any function that makes all the triangles commute. Thus

\[
(p_i \circ f)(b) = (p_i \circ g)(b)
\]
for each index $i$ and element $b \in B$. This gives

$$f(b)(i) = g(b)(i)$$

for each $b \in B$ and each index $i$. But now

$$f(b) = g(b)$$

for each $b$, so that $f = g$. ■

2.3 (a)(i) $\implies$ (ii).

Suppose that $f$ is $\epsilon - \delta$ continuous at $a$, that is

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x \in \mathbb{R})[|x - a| < \delta \implies |f(x) - f(a)| < \epsilon]$$

holds. Consider any sequence

$$x(\cdot) : \mathbb{N} \longrightarrow \mathbb{R}$$

which converges to $a$, that is

$$(\forall \gamma > 0)(\exists N \in \mathbb{N})(\forall n \in \mathbb{N})[N < n \implies |x_n - a| < \gamma]$$

holds. In this conditions ‘$\gamma$’ would normally be written ‘$\epsilon$’, but here we don’t want two ‘$\epsilon$’s to be confused. For (ii) we require

$$(\forall \epsilon > 0)(\exists N \in \mathbb{N})(\forall n \in \mathbb{N})[N < n \implies |f(x_n) - f(a)| < \epsilon]$$

to hold.

Consider any $\epsilon > 0$. The first condition gives a $\delta > 0$ such that

$$|x - a| < \delta \implies |f(x) - f(a)| < \epsilon$$

for all $x \in \mathbb{R}$. With $\gamma = \delta$ the second condition gives some $N$ such that

$$N < n \implies |x_n - a| < \delta$$

for all $n \in \mathbb{N}$. But these two implications give

$$N < n \implies |x - a| < \delta \implies |f(x) - f(a)| < \epsilon$$

for the required result.

(a)(ii) $\implies$ (i). In fact, we show the contrapositive.

$\neg$(i) $\implies$ $\neg$(ii)

We assume the negation of (i) and use AC to produce a sequence

$$x(\cdot) : \mathbb{N} \longrightarrow \mathbb{R}$$

such that

$$x_n \longrightarrow a \quad f(x_n) \not\longrightarrow f(a) \quad \text{as } n \longrightarrow \infty$$

holds.
By passing ‘¬’ through the formal definition of (i), as given above we get the following.

\[(\exists \epsilon > 0)(\forall \delta > 0)(\exists x \in \mathbb{R})[(|x - a| < \delta \text{ and } |f(x) - f(a)| \geq \epsilon)]\]

We fix such an \(\epsilon > 0\) and work with this. For each \(n \in \mathbb{N}\) we may consider \(\delta = 1/(n + 1)\) to produce a set \(A_n\) of reals where

\[x \in A_n \iff |x - a| < 1/(n + 1) \text{ and } |f(x) - f(a)| \geq \epsilon\]

for each \(x \in \mathbb{R}\). The negated condition ensures that each \(A_n\) is non-empty. Thus we have an indexed family

\[(A_n \mid n \in \mathbb{N})\]

of non-empty sets. A use of AC produces a sequence

\[x_{(\cdot)} : \mathbb{N} \longrightarrow \mathbb{R}\]

such that

\[|x_n - a| < 1/(n + 1) \quad |f(x_n) - f(a)| \geq \epsilon\]

for each \(n \in \mathbb{N}\). The left hand conditions ensure that the sequence \(x_{(\cdot)}\) converges to \(a\), and the right hand condition ensures that \(f(x_{(\cdot)})\) does not converge to \(f(a)\), and hence \(f\) is not sequentially convergent.

(b) For an arbitrary subset \(B \subseteq \mathbb{R}\) the interior \(B^o\) is the union of all the open intervals that are included in \(B\). Thus

\[a \in B^o \iff (\exists \epsilon > 0)[(a - \epsilon, a + \epsilon) \subseteq B]\]

gives this set. This is always included in \(B\).

For an arbitrary subset \(A \subseteq \mathbb{R}\) the closure \(A^-\) is the complement of the interior of the complement of \(A\). Thus

\[a \in A^- \iff \neg(\exists \epsilon > 0)[(a - \epsilon, a + \epsilon) \subseteq A']\]

gives this set (where \(A'\) is the complement of \(A\)). Let’s simplify the right hand condition. We have

\[a \in A^- \iff (\forall \epsilon > 0)[(a - \epsilon, a + \epsilon) \not\subseteq A']\]
\[\iff (\forall \epsilon > 0)(\exists x)[a - \epsilon < x < a + \epsilon \text{ and } x \notin A']\]
\[\iff (\forall \epsilon > 0)(\exists x \in A)[a - \epsilon < x < a + \epsilon]\]
\[\iff (\forall \epsilon > 0)[A \text{ meets } (a - \epsilon, a + \epsilon)]\]

which gives a neater description.

Observe that \(A \subseteq A^-\) always holds.

(b)(i) \implies (ii). Suppose (i), that is \(A = A^-\). Consider any sequence \(x_{(\cdot)}\) of elements of \(A\) and suppose \(x_n \longrightarrow a\). We require \(a \in A\), but by (i) we see that \(a \in A^-\) will suffice. Consider any \(\epsilon > 0\). By the convergence of \(x_{(\cdot)}\) we have

\[|x_n - a| < \epsilon \quad \text{that is} \quad x_n \in A \cap (a - \epsilon, a + \epsilon)\]
for all large \( n \). Since there is at least one large \( n \), we have \( a \in A^- = A \), as required.

(b)(ii) \( \Rightarrow \) (i). Suppose (ii) and consider any \( a \in A^- \). We require \( a \in A \). By the above description on \( A^- \) we have a set

\[
A_n = A \cap \left( \frac{a-1}{n+1}, \frac{a+1}{n+1} \right) \neq \emptyset
\]

for each \( n \in \mathbb{N} \). Thus we have an indexed family

\[
(A_n \mid n \in \mathbb{N})
\]

of non-empty sets. A use of AC produces a sequence

\[
x(\cdot) : \mathbb{N} \rightarrow A \quad \text{such that} \quad |x_n - a| < 1/(n+1)
\]

for each \( n \in \mathbb{N} \). This condition show that \( x_n \rightarrow a \), and hence (ii) ensures that \( a \in A \).

(c)(i) \( \Rightarrow \) (ii). Assuming (i) consider any sequence \( x(\cdot) \) of elements of \( A \). Since \( A \) is bounded above, for each \( m \in \mathbb{N} \) we may set

\[
y_m = \sup \{ x_n \mid m \leq n \}
\]

to obtain a second sequence (which may go outside of \( A \)). Observe that

\[
x_m \leq y_m \quad y_{m+1} \leq y_m
\]

for each \( m \in \mathbb{N} \). Since \( A \) is bounded below, the sequence \( y(\cdot) \) is descending and bounded below. Thus, by the completeness of \( \mathbb{R} \), this sequence has a limit \( y \), say.

For each \( n \in \mathbb{N} \) let \( f(n) \) be the least \( m > n \) such that

\[
y - 1/(n+1) < x_m \leq y < y + 1/(n+1)
\]

holds. (This does not require choice, but you should check that there is such an \( m \).) The required sub-sequence is \( n \mapsto x_{\phi(n)} \) where

\[
\phi(n) = f^{n+1}(0)
\]

for each \( n \).

(c)(ii) \( \Rightarrow \) (i). Suppose (ii), and by way of contradiction suppose \( A \) is not bounded above. Then for each \( n \in \mathbb{N} \) the set

\[
A_n = A \cap [n, \infty)
\]

is non-empty. A use of AC now produces a sequence

\[
x(\cdot) : \mathbb{N} \rightarrow A
\]

which is not bounded above and hence not convergent.

A similar argument shows that \( A \) is bounded below.

If you have a couple of hours to spare whilst going through passport control, perhaps you could sort out exactly which properties of \( \mathbb{R} \) are needed for these arguments. ■
The axiom of choice defined: Problems

2.4 If you look at the Solution 2.3 you will see that each time AC is required it is used to produce a sequence

\[ x(\cdot) : \mathbb{N} \to \mathbb{R} \]

sometimes with a more specific target. Each of these is a use of AC\(\omega\). ■

2.5 (a) The implication \((\beta) \implies (\alpha)\) is trivial.

Assuming given a function \(f\) by \((\beta)\), consider any \(a \in A\), and set

\[ \phi(n) = f^n(a) \]

for each \(n \in \mathbb{N}\).

(b) Assuming \((\alpha)\), for each \(x \in A\) consider the subset \(A_x \subseteq A\) given by

\[ y \in A_x \iff xRy \]

for \(y \in A\). This produces an \(A\) indexed family

\[ \mathcal{A} = \left\{ A_x \mid x \in A \right\} \]

of non-empty sets. By AC this family has a choice function

\[ f : A \to \bigcup \mathcal{A} \]

and this ensures that \((\beta)\) holds.

For the converse, that the given implication implies AC, we must make sure we do not use AC in the argument.

Consider any indexed family

\[ \mathcal{A} = \left\{ A_i \mid i \in I \right\} \]

of non-empty sets. We require a choice function for this family. Let

\[ A = \bigcup \mathcal{A} \times I \]

and consider the relation \(R\) on \(A\) given by

\[ (a, i)R(b, j) \iff b \in A_i \]

for each \(a, b \in \bigcup \mathcal{A}\) and \(i, j \in I\). Since each \(A_i\) is non-empty we see that \((\alpha)\) holds for this pair \(A, R\). Thus, by the given implication, we have \((\beta)\), that is there is a function \(f : A \to A\) with

\[ (a, i)Rf(a, i) \]

for each element \(a \in \bigcup \mathcal{A}\) and index \(i \in I\). Since the target of \(f\) is a product, we have two functions

\[ l : A \to \bigcup \mathcal{A} \quad r : A \to I \]

such that

\[ f(a, i) = (l(a, i), r(a, i)) \]
for each element \(a \in \bigcup \mathcal{A}\) and index \(i \in I\). In particular, we have
\[
l(a, i) \in A_i
\]
for each element \(a \in \bigcup \mathcal{A}\) and index \(i \in I\). We can now fix any \(a \in \bigcup \mathcal{A}\), and then \(i \mapsto l(a, i)\) is a choice function for the given indexed family.

(c) Assuming AC parts (b, a) give
\[
(\alpha) \implies (\beta) \implies (\gamma)
\]
which gives AC \(\Rightarrow\) DC.

The implication DC \(\Rightarrow\) AC\(_\omega\) needs just a little bit more work. Consider any \(\mathbb{N}\)-indexed family
\[
\mathcal{A} = (A_n \mid n \in \mathbb{N})
\]
of non-empty sets. We require a choice function for this family, and to obtain that we may apply DC. Let
\[
A = \bigcup \mathcal{A} \times \mathbb{N}
\]
and consider the relation \(R\) on \(A\) given by
\[
(a, m) R (b, n) \iff b \in A_m \text{ and } n = m + 1
\]
for each \(a, b \in \bigcup \mathcal{A}\) and \(m, n \in \mathbb{N}\). Since each \(A_m\) is non-empty we see that \((\alpha)\) holds for this pair \(A, R\). Thus, DC gives a function
\[
\phi : \mathbb{N} \longrightarrow A \text{ where } \phi(k) R \phi(k + 1)
\]
for each \(k \in \mathbb{N}\). Observe that if
\[
\phi(k) = (a, m) \text{ then } \phi(k + 1) = (b, m + 1)
\]
for some \(b \in A_m\).

Now consider
\[
\phi(0) = (a, m)
\]
for some \(a \in \bigcup \mathcal{A}\) and \(m \in \mathbb{N}\). We then have
\[
\phi(1) = (b, m + 1) \text{ where } b \in A_m
\]
for some \(b\). More generally, for each \(k \in \mathbb{N}\) we have
\[
\phi(k + 1) = (c, m + k + 1) \text{ where } c \in A_{m+k}
\]
for some \(c\). Thus \(\phi\) provides a choice function for
\[
(A_n \mid n \in \mathbb{N}, m \leq n)
\]
which is most of the given indexed family.

By Exercise 2.1 we can produce a member of
\[
A_0 \times \cdots \times A_{m-1}
\]
without using any kind of choice. Thus we obtain a choice function for the whole of the
given family.

2.6 (a) We require
\[(\alpha \star \beta) \star \gamma = (\alpha \star (\beta \star \gamma))\]
for each choice functions \(\alpha, \beta, \gamma \in A\). To verify this we evaluate each side at an arbitrary
index \(i\). Thus we have
\[
((\alpha \star \beta) \star \gamma)(i) = (\alpha \star \beta)(i) \star \gamma(i) = (\alpha(i) \star i \beta(i)) \star i \gamma(i)
\]
and then the associativity of \(\star_i\) gives the required result.

We also require
\[(\alpha \star 1) = \alpha = (1 \star \alpha)\]
but these follow by a similar, but easier, calculation.

(b) For a given index \(i\) we require
\[p_i(\alpha \star \beta) = p_i(\alpha) \star_i p_i(\beta)\]
for each \(\alpha, \beta \in A\). But, by definition of \(p_i\), we have
\[p_i(\alpha \star \beta) = (\alpha \star \beta)(i) \quad p_i(\alpha) \star_i p_i(\beta) = \alpha(i) \star_i \beta(i)\]
so the required result follows by the definition of \(\star\).

We also require
\[p_i(1) = i\]
but this follow by am almost trivial calculation.

(c) By Exercise 2.2 there is only one possible function \(f\), that given by
\[f(b)(i) = f_i(b)\]
for each \(b \in B\) and index \(i\). Thus it suffices to show that this function
\[f : B \longrightarrow A\]
is a morphism.

We require
\[f(b \star c) = f(b) \star f(c)\]
for each \(b, c \in B\). Here \(\star\) is the operation in \(\mathfrak{B}\). To verify this we evaluate both sides at
an arbitrary index \(i\). For the left hand side we have
\[f(b \star c)(i) = f_i(b \star c) = f_i(b) \star_i f_i(c)\]
where the second step holds because \(f_i\) is a given morphism. For the right hand side we have
\[(f(b) \star f(c))(i) = f(b)(i) \star_i f(c)(i) = f_i(b) \star_i f_i(c)\]
where the first step holds by the construction of \(\star\).

The preservation of the unit is almost immediate.  ■