Analysis 1, Solutions to problem set 4

Problem 1 (Cardinality)

(a) Let us use the letter $\mathcal{F}$ to denote the set of all functions $f : \{0,1\} \to \mathbb{N}$.

An element $f \in \mathcal{F}$ is completely determined by the values it takes at 0 and at 1, i.e. by looking at the natural number $f(0)$ and *then* at the natural number $f(1)$. This information can be encoded in the "ordered pair" $(f(0), f(1))$.

The following picture illustrates a map, denoted Φ, from the set natural numbers $\mathbb{N}$ onto the set $\mathcal{F}$:

$$\begin{array}{ccccccc}
(0,0) & (0,1) & (0,2) & (0,3) & (0,4) & (0,5) & \cdots \\
(1,0) & (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & \cdots \\
(2,0) & (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & \cdots \\
(3,0) & (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & \cdots \\
(4,0) & (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots &
\end{array}$$

which should be interpreted as follows:

$$\begin{aligned}
0 &\overset{\Phi}{\mapsto} (f_0 : \{0,1\} \to \mathbb{N},\ f_0(0) = 0, f_0(1) = 0) \\
1 &\overset{\Phi}{\mapsto} (f_1 : \{0,1\} \to \mathbb{N},\ f_1(0) = 0, f_1(1) = 1) \\
2 &\overset{\Phi}{\mapsto} (f_2 : \{0,1\} \to \mathbb{N},\ f_2(0) = 1, f_2(1) = 0) \\
3 &\overset{\Phi}{\mapsto} (f_3 : \{0,1\} \to \mathbb{N},\ f_3(0) = 2, f_3(1) = 0) \\
4 &\overset{\Phi}{\mapsto} (f_4 : \{0,1\} \to \mathbb{N},\ f_4(0) = 1, f_4(1) = 1) \\
&\vdots
\end{aligned}$$

By construction, it is clear that the map $\Phi : \mathbb{N} \to \mathcal{F}$ just described is well-defined and bijective. Theorem 1.33 implies in particular that there exists an injective map $\Psi : \mathcal{F} \to \mathbb{N}$. It follows that the set $\mathcal{F}$ is countable, as desired.

Remark. The set $\mathcal{F}$ is usually denoted in the literature by "$\mathbb{N} \times \mathbb{N}$" to emphasize the fact that it consists of "ordered pairs" of natural numbers. We will obey this tradition in the problems below.

(b) Let $\mathbb{Q}_+$ denote the set of positive rationals, i.e.

$$\mathbb{Q}_+ := \{q \in \mathbb{Q} : q > 0\}.$$

For every $q \in \mathbb{Q}_+$, there exists at least one pair $(n,m) \in \mathbb{N} \times \mathbb{N}$ such that $q = \frac{m}{n}$. Therefore we can find an injection $\iota : \mathbb{Q}_+ \to \mathbb{N} \times \mathbb{N}, q \mapsto (n,m)$. In part (a) we proved that the set $\mathbb{N} \times \mathbb{N}$ is countable, i.e. there exists an injection $\phi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$. Since the composition of injective functions is injective, it follows that the function $\phi \circ \iota$ provides an injection from $\mathbb{Q}_+$ to $\mathbb{N}$. It follows that $\mathbb{Q}_+$ is countable.

In a similar way, the set $\mathbb{Q}_-$ of negative rationals is countable. It follows from Problem 0 of this week's Präsenzblatt that the set

$$\mathbb{Q} = \mathbb{Q}_- \cup \{0\} \cup \mathbb{Q}_+$$

is countable, as desired.

(c) Aiming at a contradiction, let $\varphi : \mathbb{N} \to \mathbb{Q}$ be an order-preserving bijection from the set of natural numbers to the set of rationals. Consider the rational numbers $\varphi(0)$ and $\varphi(1)$. Since φ is in particular an injection, it follows that $\varphi(0) \neq \varphi(1)$. Since $0 < 1$, we actually have that $\varphi(0) < \varphi(1)$ since φ is order-preserving by assumption. Consider the midpoint q of the rational numbers $\varphi(0)$ and $\varphi(1)$,

$$q := \frac{\varphi(0) + \varphi(1)}{2}.$$

This is still a rational number, which satisfies $\varphi(0) < q < \varphi(1)$. We claim that $\varphi(n) \neq q$ for every $n \in \mathbb{N}$. We already know that $\varphi(0) \neq q$ and that $\varphi(1) \neq q$. Any other natural number $n \in \mathbb{N} \setminus \{0, 1\}$ satisfies $1 < n$, and so $\varphi(1) < \varphi(n)$ by the order-preserving property of φ. It follows that

$$q < \varphi(1) < \varphi(n),$$

and so $\varphi(n) \neq q$, as claimed. Thus φ does not surject onto $\mathbb{Q}$, and as such it cannot be a bijection. The contradiction resulted from assuming the existence of an order-preserving bijection from the set of natural numbers to the set of rationals. Thus no such order-preserving bijection exists, as we wanted to prove.

Problem 2 (Recursion)

(a) Let us start by using induction on m to prove that, for every $m \in \mathbb{N}$,

$$\forall n \in \mathbb{N} : \nu(n) + m = \nu(n + m), \tag{1}$$

which, using the definition of addition $+$, can be stated in the following equivalent form:

$$\forall n \in \mathbb{N} : \nu^m(\nu(n)) = \nu(\nu^m(n)).$$

The base case $m = 0$ follows immediately from the first property of addition established in Theorem 1.24. Indeed, for any $n \in \mathbb{N}$,

$$\nu^0(\nu(n)) = \nu(n) + 0 = \nu(n) = \nu(n + 0) = \nu(\nu^0(n)).$$

Assuming the induction hypothesis

$$\forall n \in \mathbb{N} : \nu^m(\nu(n)) = \nu(\nu^m(n)), \tag{2}$$

let us prove that

$$\forall n \in \mathbb{N} : \nu^{\nu(m)}(\nu(n)) = \nu(\nu^{\nu(m)}(n)).$$

With this purpose in mind, let $n \in \mathbb{N}$ be arbitrary. Then:

$$\nu^{\nu(m)}(\nu(n)) \overset{(i)}{=} \nu(\nu^m(\nu(n))) \overset{(ii)}{=} \nu(\nu(\nu^m(n))) \overset{(iii)}{=} \nu(\nu^{\nu(m)}(n)),$$

as desired. Here, (i) and (iii) are a consequence of property 2 of Theorem 1.22, whereas (ii) amounts to the induction hypothesis (2). Thus (1) is established for every $m \in \mathbb{N}$.

Let us now use induction on n to show that, for every $n \in \mathbb{N}$ and every $p \in \mathrm{Dom}(g)$,

$$\forall m \in \mathbb{N} : g^{n+m}(p) = g^n(g^m(p)).$$

The base case $n = 0$ is again a consequence of the fact that 0 is the neutral element for addition. In fact, for any $m \in \mathbb{N}$,

$$g^{0+m}(p) = g^m(p) = g^0(g^m(p)).$$

Assuming the induction hypothesis

$$\forall m \in \mathbb{N} : g^{n+m}(p) = g^n(g^m(p)), \tag{3}$$

let us prove that

$$\forall m \in \mathbb{N} : g^{\nu(n)+m}(p) = g^{\nu(n)}(g^m(p)).$$

With that purpose in mind, let $m \in \mathbb{N}$ be arbitrary. Then

$$g^{\nu(n)+m}(p) \overset{(i)}{=} g^{\nu(n+m)}(p) \overset{(ii)}{=} g(g^{n+m}(p)) \overset{(iii)}{=} g(g^n(g^m(p))) \overset{(iv)}{=} g^{\nu(n)}(g^m(p)),$$

as desired. Here, (i) is a consequence of property (1) proved above and (iii) follows from the induction hypothesis (3). On the other hand, steps (ii) and (iv) are a consequence of property 2 of Theorem 1.22.

More precisely: since g satisfies the hypothesis of Theorem 1.22 and $p \in \mathrm{Dom}(g)$, there exists a sequence $h : \mathbb{N} \to \mathrm{Dom}(g)$ such that $h(0) = p$ and

$$\forall x \in \mathbb{N} : h(\nu(x)) = g(h(x)). \quad (*)$$

In particular, since (by definition) $h(k) = g^k(p)$ for every $k \in \mathbb{N}$,

$$g^{\nu(n+m)}(p) = h(\nu(n+m)) \overset{(*)}{=} g(h(n+m)) = g(g^{n+m}(p)).$$

This establishes (ii), and a similar argument establishes (iv). This concludes the proof.

(b) Part (a) with $g = \nu$ and $p = k \in \mathbb{N} = \mathrm{Dom}(\nu)$ tells us that, for every natural numbers n, m,

$$\nu^{n+m}(k) = \nu^n(\nu^m(k)). \tag{4}$$

Then

$$\begin{aligned} k + (n+m) &= \nu^{n+m}(k) \quad \text{(by definition of +)} \\ &= \nu^n(\nu^m(k)) \quad \text{(by (4))} \\ &= \nu^n(k+m) \quad \text{(by definition of +)} \\ &= (k+m) + n. \quad \text{(by definition of +)} \end{aligned} \tag{5}$$

The result follows from this and the commutativity property (C) of addition which was already proved in class. Indeed,

$$(n+m) + k \overset{(C)}{=} k + (n+m) \overset{(5)}{=} (k+m) + n \overset{(C)}{=} n + (k+m) \overset{(C)}{=} n + (m+k).$$

Problem 3 (Surjectivity and injectivity)

(a) From the definition of surjectivity, the hypothesis implies that

$$\forall y \in Y \exists x \in X : f(x) = y.$$

Applying the axiom of choice (Rule 34) to this statement and the identity function on Y, denoted I_Y, we conclude the existence of a function g for which

$$\forall y : (f(g(y)) = y \wedge g(y) \neq g \wedge I_Y(y) \neq I_Y) \vee (g(y) = g \wedge I_Y(y) = I_Y). \tag{6}$$

If $y \in Y$, then $I_Y(y) \neq I_Y$, and so $g(y) \neq g$ (and $f(g(y)) = y$). Conversely, if $g(y) \neq g$, then $I_Y(y) \neq I_Y$ and so $y \in Y$. This shows that $\mathrm{Dom}(g) = Y$. Also, if $x \in \mathrm{Ran}(g)$, then

$$\exists y : g(y) = x \wedge x \neq g.$$

In particular, $f(x) = f(g(y)) = y$ because $g(y) \neq g$. Since $y \neq f$ (because $y \in \mathrm{Dom}(g) = Y = \mathrm{Ran}(f)$), it follows that $x \in \mathrm{Dom}(f) = X$. This shows that $\mathrm{Ran}(g) \subset X$.

Hence we will be done once we show that g is injective, i.e.,

$$\forall y_1 \in Y \forall y_2 \in Y : g(y_1) \neq g(y_2) \vee y_1 = y_2.$$

Let $y_1, y_2 \in Y$ be such that $g(y_1) = g(y_2)$. We want to show that $y_1 = y_2$. Since $I_Y(y_1) \neq I_Y$ and $I_Y(y_2) \neq I_Y$, it follows from (6) that $f(g(y_1)) = y_1$ and that $f(g(y_2)) = y_2$. But then

$$y_1 = f(g(y_1)) = f(g(y_2)) = y_2,$$

as desired.

(b) By definition of injectivity,

$$\forall x_1 \in X \forall x_2 \in X : f(x_1) \neq f(x_2) \vee x_1 = x_2.$$

In particular, given $y \in \mathrm{Ran}(f) \subset Y$, there exists exactly one $x \in X$ such that $f(x) = y$. Define g on such y to be equal to that specific x, i.e. $g(y) = x$. On the other hand, if $y \in Y \setminus \mathrm{Ran}(f)$, pick a fixed $x_0 \in X$ and define $g(y) = x_0$ (this can be done because without loss of generality $X \neq \emptyset$). The function g thus defined: for $y \in Y$,

$$g(y) = \begin{cases} x & \text{if } y \in \mathrm{Ran}(f) \text{ and } f(x) = y \\ x_0 & \text{if } y \in Y \setminus \mathrm{Ran}(f), \end{cases}$$

and $g(y) = g$ otherwise, satisfies:

(i) $\mathrm{Dom}(g) = Y$,

(ii) $\mathrm{Ran}(g) \subset X$,

(iii) g is surjective.

Parts (i) and (ii) follow directly from the way the function g was constructed. To check (iii), we additionally have to verify that $X \subset \mathrm{Ran}(g)$. But this is immediate since $g(f(x)) = x$ for every $x \in X$.

Problem 4 (Arithmetic Mean-Geometric Mean inequality)

Proof 1. This proof uses only elementary arithmetic rules and mathematical induction.
The base case $n = 1$ is trivial to verify since $x_1 = (x_1/1)^1$. Let us assume that the statement has been verified for all choices of n nonnegative real numbers. Consider $n + 1$ nonnegative real numbers $x_1, x_2, \ldots, x_n, x_{n+1}$ with arithmetic mean A defined via

$$(n+1)A := x_1 + x_2 + \ldots + x_n + x_{n+1}. \tag{7}$$

If all the numbers x_i equal A, we are done. Otherwise we can find one number which is strictly larger than A and one number which is strictly smaller than A, say $x_n > A$ and $x_{n+1} < A$. In particular, the real numbers $x_n - A$ and $A - x_{n+1}$ are (strictly) positive, and so

$$(x_n - A)(A - x_{n+1}) > 0. \tag{8}$$

Let us now consider the n numbers $x_1, x_2, \ldots, x_{n-1}, x$, where $x := x_n + x_{n+1} - A$. Note that x is a positive real number. Indeed, since $x_{n+1} \geq 0$,

$$x = x_n + x_{n+1} - A \geq x_n - A > 0.$$

The crucial observation is that A *is still the arithmetic mean of the* n *numbers* $x_1, x_2, \ldots, x_{n-1}, x$. Indeed, from (7) it follows at once that

$$nA = x_1 + x_2 + \ldots + x_{n-1} + \underbrace{x_n + x_{n+1} - A}_{=x}.$$

Using the induction hypothesis, we thus conclude that $A^n \geq x_1 x_2 \cdots x_{n-1} x$, and so

$$A^{n+1} = A^n \cdot A \geq (x_1 x_2 \cdots x_{n-1} x) A. \tag{9}$$

Now,

$$\begin{aligned} xA - x_n x_{n+1} &= (x_n + x_{n+1} - A)A - x_n x_{n+1} \quad \text{(by definition of } x) \\ &= (x_n - A)(A - x_{n+1}) \quad \text{(expand brackets)} \\ &> 0, \quad \text{(by (8))} \end{aligned}$$

and so

$$xA > x_n x_{n+1} \geq 0, \tag{10}$$

and thus $A > 0$. Thus, if at least one of the numbers $x_1, x_2, \ldots, x_{n-1}$ is zero, then we already have strict inequality in (9). Otherwise the right-hand side of (9) is positive and *strict* inequality can be derived from (9) and (10):

$$\begin{aligned} A^{n+1} &\geq (x_1 x_2 \cdots x_{n-1} x) A \\ &= x_1 x_2 \cdots x_{n-1} (xA) \\ &> x_1 x_2 \cdots x_{n-1} (x_n x_{n+1}) \\ &= x_1 x_2 \cdots x_{n+1}. \end{aligned}$$

In particular, the inequality is an equality if and only if all the x_i are the same.

Proof 2. (Sketch) An alternative elegant proof due to Cauchy is available and involves a non-standard kind of induction which sometimes goes by the name of "forward-backward induction". To briefly describe it, let $P(n)$ denote the statement of the inequality we want to prove,

$$x_1 \cdot x_2 \cdots x_n \leq \Big(\frac{x_1 + x_2 + \ldots + x_n}{n}\Big)^n.$$

For $n = 2$, we have that

$$x_1 \cdot x_2 \leq \Big(\frac{x_1 + x_2}{2}\Big)^2 \text{ if and only if } (x_1 - x_2)^2 \geq 0,$$

which is true. Then we proceed in the following two steps, which will clearly imply the full result:

(i) $P(n) \Rightarrow P(n-1)$;

(ii) $P(n)$ and $P(2) \Rightarrow P(n-1)$.

For part (i), let $A := \frac{1}{n-1} \sum_{k=1}^{n-1} x_k$, and observe that $P(n)$ implies

$$(x_1 \cdot x_2 \cdots x_{n-1}) A \leq \Big(\frac{\sum_{k=1}^{n-1} x_k + A}{n}\Big)^n = \Big(\frac{(n-1)A + A}{n}\Big)^n = A^n,$$

and hence

$$x_1 \cdot x_2 \cdots x_{n-1} \leq A^{n-1} = \Big(\frac{x_1 + x_2 + \ldots + x_{n-1}}{n-1}\Big)^{n-1},$$

which is $P(n-1)$.
Part (ii) is left to the reader.