KMB, 12/12/16

M1F Foundations of Analysis, Problem Sheet 10.

1*. Did you manage this one?

2.

(i) f is bijective; indeed if we define $g = f$ then g is a two-sided inverse function for f. For $f(f(0)) = f(0) = 0$, and if $x \neq 0$ then $f(f(x)) = f(1/x) = 1/(1/x) = x$, so $f \circ f$ is the identity function $\mathbf{R} \to \mathbf{R}$.

(ii) $f : \mathbf{Z} \to \mathbf{Z}$, $f(n) = 2n + 1$. This is injective, because if $f(a) = f(b)$ then $2a + 1 = 2b + 1$ and hence $2a = 2b$, so $a = b$. But it is not surjective, as $f(n)$ is always odd, so there cannot be any n such that $f(n) = 2$ (indeed if $f(n) = 2$ then $2n = 1$ but no integer satisfies this equation).

(iii) $f : \mathbf{R} \to \mathbf{R}$, $f(x) = x^3$. This is bijective. It's pretty obvious that if $g(x) = x^{1/3}$ then g is a two-sided inverse for f, but if I was going to be super-fussy I would say that in the question I only said that you could assume that every positive real has a unique positive real cube root, so first you should make the following observations. (a) The cube of a non-positive real is non-positive, hence every positive real has a unique real cube root; (b) the cube of a non-zero number is non-zero, so zero has a unique real cube root (namely zero); (c) if $y^3 = x$ then $(-y)^3 = -x$, from which it follows that every negative number has a unique real cube root. We've just checked carefully that every real number x has a unique real cube root, and if we define $g(x)$ to be the unique real cube root of x then $g(y)^3 = y$ and $g(x^3) = x$ (by uniqueness), hence $f(g(y)) = y$ and $g(f(x)) = x$ for all x, y meaning that f and g are inverse functions.

(iv) $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3$ if the Riemann hypothesis is true, and $f(x) = -x$ if not. This function is bijective, and the two-sided inverse function is g defined by $g(y) = y^{1/3}$ if the Riemann hypothesis is true, and $g(y) = -y$ if it is false. A case by case check shows that whether or not the Riemann Hypothesis is true, $f \circ g$ and $g \circ f$ are both the identity function.

(v) This question is easier than I thought – one of my tutees pointed out that $f(n+1) - f(n)$ was a quadratic with negative discriminant, hence (by completing the square) $f(n+1) > f(n)$ for all n. This can be used to prove both that f is injective ($a < b$ implies $f(a) < f(b)$ by induction on $b - a$) and not surjective ($f(1) = 0$ and $f(2) = 3$ so there can be no integer n such that $f(n) = 1$). Here is my original answer though (much more complicated).

This function is not surjective. For example I claim there can be no integer n such that $f(n) = 1$; indeed such an integer n would satisfy $n^3 - 2n^2 + 2n = 2$ and hence $n(n^2 - 2n + 2) = 2$; hence n would have to be a divisor of 2. But the only divisors of 2 are ± 1 and ± 2, and $f(1) = 0$, $f(2) = 3$, $f(-1) = -6$ and $f(-2) = -21$, so none of these work.

It is injective however, although this is perhaps a little tough to prove. We do it by contradiction. Say $m, n \in \mathbf{Z}$ with $m \neq n$ and $f(m) = f(n)$. Then $f(m) - f(n) = 0$, so $(m^3 - n^3) - 2(m^2 - n^2) + 2(m - n) = 0$. Because $m \neq n$ we can divide out by $m - n$ and deduce $m^2 + mn + n^2 - 2(m + n) + 2 = 0$ and our job is to show that this equation has no solutions. By completing the square and then multiplying by 12 to clear denominators we deduce that $3(2m + n - 2)^2 + (3n - 2)^2 + 8 = 0$ and this has no real solutions, let alone integer ones, so this is the contradiction we seek.

3.

(i) f is not defined at zero, so it is not a function with domain $\mathbf{R}$.

(ii) f is not defined for $x < 0$ as whatever you mean by $\sqrt{x}$ it can't be real.

(iii) $f(0) = 1/2$ which is not in the codomain.

(iv) We don't say which solution (and sometimes there is more than one – for example $y^3 - y = 0$ has three real solutions). If we were careful to explain exactly which solution we chose (for example we could choose the largest real solution) then this would be well-defined (but it would not be continuous – can you find an example of a discontinuity if we chose the largest real solution in every case?)

(v) If $|x| < 1$ then $1 + x + x^2 + x^3 + \cdots = 1/(1 - x)$ (you will see a proof of this next term when you will also learn rigorously what the definition of an infinite sum is). However if $|x| > 1$

then the sum does not converge (even though $1/(1-x)$ makes perfect sense) so as it stands this function is not defined when $|x| > 1$.

3.

(i) Choose $y \in Y$ and set $x = g(y)$. We are given $f \circ g$ is the identity function, and this implies $(f \circ g)(y) = y$, so $f(g(y)) = y$ so $f(x) = y$. Hence y is in the image of f. But $y \in Y$ was arbitrary, hence f is surjective.

(ii) Say $x_1, x_2 \in X$ and $f(x_1) = f(x_2)$. Applying g we deduce that $g(f(x_1)) = g(f(x_2))$. But $g \circ f$ is the identity function $X \to X$, so $(g \circ f)(x) = x$ for all $x \in X$. In particular $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$, and hence f is injective.

4*.

(i) $(f \circ g)(x) = f(g(x)) = f(2x) = (2x)^2 + 3 = 4x^2 + 3$.
(ii) $(g \circ f)(x) = g(f(x)) = g(x^2 + 3) = 2x^2 + 6$.
(iii) The function sends x to $f(x)g(x) = (x^2 + 3)(2x) = 2x^3 + 6x$.
(iv) The function sends x to $f(x) + g(x) = x^2 + 3 + 2x = x^2 + 2x + 3$.
(v) $x \mapsto f(g(x))$ is the same function as $x \mapsto (f \circ g)(x)$ so the answer is the same as (i).

5. Say $A = \{x_1, x_2, \ldots, x_a\}$. To define a function $A \to B$ all we have to do is to say exactly where each element of A gets sent, and each element must be sent to an element of B, so there are b choices for where to send each x_i. The total number of ways we can do this is hence $b \times b \times \cdots \times b$, one "$b$" for each element of A. But there are a elements of A, so the total number of functions is b^a.

This argument seems to be evidence to support the convention that $0^0 = 1$, because there is exactly one function from the empty set to itself, namely the empty function (corresponding to the empty subset of $\emptyset \times \emptyset$). But in general 0^0 is undefined, by convention, and what is definitely clear is that there is no way to define it in such a way to make it a continuous function of each variable, because 0^r is probably most sensibly defined to be zero if $r > 0$. In fact, in practice, people only talk about x^y in the following situations:

(i) If $x \in \mathbf{R}_{>0}$ and $y \in \mathbf{C}$ then x^y can be defined as $e^{y \log(x)}$ and this satisfies all the nice properties that you want.

(ii) If $x \in \mathbf{C}$ with $x \neq 0$ and $y \in \mathbf{Z}_{\geq 0}$ then x^y can be defined in the usual way: if $y > 0$ then it's the product of y copies of x (or do it by induction if you want to be a formalist), if $y = 0$ then it's 1 and if $y < 0$ then it's $1/x^{-y}$. Again all the standard rules for powers work here.

(iii) If $x, y \in \mathbf{C}$ then choose a branch of logarithm on the complexes and again define x^y as $e^{y \log(x)}$. The problem with this definition is that it is not continuous where you made the cut for log, and does not satisfy the usual rules for powers like $(x_1)^y (x_2)^y = (x_1 x_2)^y$, because once you make a branch cut to define log it's not true that $\log(x_1) + \log(x_2) = \log(x_1 x_2)$. Note that this problem does not occur if we just restrict to x_1, x_2 positive reals.

6.

(i) We know that if $y \in Y$ then the set $\{x \in X : f(x) = y\}$ is non-empty, so we just choose a random element of it once and for all, and define $g(y)$ to be this element. Can we build such a function this way? Sure (although (formalist hat on) to guarantee that the resulting subset of $X \times Y$ is actually a set(!) we need to invoke the axiom of choice (moral: invoke the axiom of choice) (formalist hat off)). This definition of g is not at all "natural" though, in the sense that if we had an explicit function f and you chose a function g as above and called it g_1, and your friend chose another example and called it g_2, then in general the chances that you and your friend had chosen the same function would be very small.

(ii) Funnily enough such a function does not exist, because X could be the empty set and Y could be a non-empty set; then the empty function $f : X \to Y$ is injective and there are no functions $Y \to X$ at all. However if X is non-empty then it's OK; choose $x_0 \in X$, and define $g : Y \to X$ by $g(y) = x$ if y is in the image of f and $f(x) = y$ (such an x is unique) and $g(y) = x_0$ otherwise.

(iii) The argument is easiest to explain if $X \cap Y$ is empty. Then we can define Z to be the union of X and Y. Now X is a subset of Z so there's a natural injection $g : X \to Z$. Furthermore we can define $h : Z \to Y$ by $h(z) = f(z)$ if $z \in X$ and $h(z) = z$ if $z \in Y$. Then h is clearly surjective because if $y \in Y$ then $y \in Z$ and $h(y) = y$. Furthermore it is easy to check that $f = h \circ g$.

If $X \cap Y$ is not empty then we replace Y by any set Y' for which there is a bijection $i : Y' \to Y$ and such that $Y' \cap X$ is empty. Such a set Y' does exist but one would have to carefully read the axioms of mathematics to verify this (as far as I know – is there a simple trick? I want to define $Y' = \{X\} \times Y$ but to check that this is definitely disjoint from X I would have to explicitly explain how to define ordered pairs in set theory and then invoke the axiom of foundation. This situation can't be completely trivial because if there existed a set of all sets and X were this set then it would not be true!). Once you're satisfied that Y' exists then we can let Z be $X \cup Y'$ and define h by $h(x) = f(x)$ and $h(y') = i(y')$ and the same proof works.

The reason that this all feels a bit weird is that this construction is not at all natural. For example Y' can actually be replaced by any set disjoint from X and equipped with a surjection to Y, so again the construction is not natural, although the example I give is probably "universal" in some sense which you will learn about much later on if you do some of the algebra courses in your third year.

7. To check that $(f \circ g) \circ h = f \circ (g \circ h)$ we first observe that all the instances of $\circ$ actually make sense (for example $f : C \to D$ and $g : B \to C$, so $(f \circ g)$ is a well-defined function $B \to D$ etc) and as an outcome of this we see that $(f \circ g) \circ h$ and $f \circ (g \circ h)$ are both functions $A \to D$ and in particular it makes sense to ask if they are equal.

Now what does it *mean* for two functions $A \to D$ to be equal? It means that for all $a \in A$, the values of the two functions coincide at a. In other words, in our case it means that $((f \circ g) \circ h)(a) = (f \circ (g \circ h))(a)$, so this is what we need to check. However, by continually appealing to the definition of $\circ$, we see that

$$\begin{aligned} &((f \circ g) \circ h)(a) \\ &= (f \circ g)(h(a)) \\ &= f(g(h(a))) \end{aligned}$$

and

$$\begin{aligned} &(f \circ (g \circ h)(a) \\ &= f((g \circ h)(a)) \\ &= f(g(h(a))) \end{aligned}$$

so both $((f \circ g) \circ h)(a)$ and $(f \circ (g \circ h))(a)$ equal $f(g(h(a)))$, which, let's face it, are the only possible thing that they could have equalled, because there is no other conceivable way of defining an element of D given an element of A.

8.

(i) Let's count X: in other words let's write $X = \{x_1, x_2, x_3, \ldots\}$. Then Y is a subset of X, so Y looks like $\{x_3, x_{10}, x_{12345}, \ldots\}$ and what we need to do is to come up with a bijection between this and $\mathbf{N}$. But it's clear how to do such a thing – for the Y above we would set $f(1) = x_3$, $f(2) = x_{10}$ and so on, and in general Y must have the form $\{x_s : s \in S\}$ where S is an infinite subset of $\mathbf{N}$, and so we can write $S = \{s_1, s_2, s_3, \ldots\}$ and then define our bijection $\mathbf{N} \to Y$ by sending n to x_{s_n}.

(ii) If X and Y are countable then $X \cup Y$ is countable (c.f., counting $\mathbf{Z}$). The reals are not countable (part (i)) but the rationals are, so if the irrationals were also countable then then the reals would be too, a contradiction. Thus the irrationals must be uncountable. The complexes are clearly uncountable because they contain the reals which are uncountable so again we're done by (i). Finally $\mathbf{Q}(i)$ is countable, because as a set it clearly bijects with $\mathbf{Q} \times \mathbf{Q}$, and the product of two countable sets is countable,