Analysis 1
Recitation Session of Week 3

October 3, 2014

1 Homework Sheet Number 3

1.1 Inverse Image

- For a map $f : X \to Y$, and a subset $B \subset Y$, we have defined a new set, a subset of X, denoted by $f^{-1}(B)$ which is defined as $\boxed{f^{-1}(B) := \{\, x \in X \mid f(x) \in B \,\}}$.
- The inverse image possess many nice properties with set operations which are either only partially true or completely true for the image of $A \subset X$ under f: $\boxed{f(A) := \{\, y \in Y \mid \exists x \in A : f(x) = y \,\}}$.
- Here is an example:
  Claim: $f\left(\bigcup_{j\in I} A_j\right) = \bigcup_{j\in I} f(A_j)$
  Proof:
  - *Claim 1:* $f\left(\bigcup_{j\in I} A_j\right) \subseteq \bigcup_{j\in I} f(A_j)$
    Proof:
    1. Let $y_0 \in f\left(\bigcup_{j\in I} A_j\right)$ be given.
    2. $\iff y_0 \in \left\{ y \in Y \,\middle|\, \left[\exists x \in \bigcup_{j\in I} A_j\right] : f(x) = y \right\}$
    3. $\iff \left(\exists x \in \bigcup_{j\in I} A_j\right) : f(x) = y_0$
    4. $\iff (\exists j_0 \in I : \exists x \in A_{j_0}) : f(x) = y_0$.
    5. $\iff \exists j_0 \in I : (\exists x \in A_{j_0} : f(x) = y_0)$ (basic result in formal logic).
    6. $\iff \exists j_0 \in I : (y_0 \in \{\, y \in Y \mid [\exists x \in A_{j_0}] : f(x) = y \,\})$
    7. $\iff y_0 \in \bigcup_{j\in I} \{\, y \in Y \mid [\exists x \in A_j] : f(x) = y \,\}$
    8. $\iff y_0 \in \bigcup_{j\in I} f(A_j)$
    9. Thus we have shown that an arbitrary element of $f\left(\bigcup_{j\in I} A_j\right)$ is also an element of $\bigcup_{j\in I} f(A_j)$, which means that *any* element of $f\left(\bigcup_{j\in I} A_j\right)$ is also an element of $\bigcup_{j\in I} f(A_j)$.
  - *Claim 2:* $f\left(\bigcup_{j\in I} A_j\right) \supseteq \bigcup_{j\in I} f(A_j)$
    Proof:

    * Let $y_0 \in \bigcup_{j\in I} f(A_j)$ be given.
    * Because all the implications above were $\Longleftrightarrow$ (two-sided), we can use the procedure above to conclude that $y_0 \in f\left(\bigcup_{j\in I} A_j\right)$.
    * Thus we have shown that an arbitrary element of $\bigcup_{j\in I} f(A_j)$ is also an element of $f\left(\bigcup_{j\in I} A_j\right)$, which means that *any* element of $\bigcup_{j\in I} f(A_j)$ is also an element of $f\left(\bigcup_{j\in I} A_j\right)$.

■

- Here is another example:
  Claim: $B_1 \subseteq B_2 \Longrightarrow f^{-1}(B_1) \subseteq f^{-1}(B_2)$ for any two subset B_1 and B_2 of Y.
  Proof:
  - Let $x_0 \in f^{-1}(B_1)$ be given.
  - $\Longleftrightarrow x_0 \in \{\, x \in X \mid f(x) \in B_1 \,\}$
  - $\Longleftrightarrow f(x_0) \in B_1$
  - But $B_1 \subseteq B_2$ according to our assumption, so that $f(x_0) \in B_2$.
  - $f(x_0) \in B_2 \Longleftrightarrow x_0 \in \{\, x \in X \mid f(x) \in B_2 \,\} \Longleftrightarrow x_0 \in f^{-1}(B_2)$.
  - Because x_0 was arbitrary, we conclude that $f^{-1}(B_1) \subseteq f^{-1}(B_2)$.

  ■

- *Claim:* $A_1 \subseteq A_2 \Longrightarrow f(A_1) \subseteq f(A_2)$.
  Proof:
  - Let $y_0 \in f(A_1)$ be given.
  - That means that $\exists x \in A_1$ such that $f(x) = y_0$.
  - But $A_1 \subseteq A_2$, so that $x \in A_2$.
  - That means that $\exists x \in A_2$ such that $f(x) = y_0$.
  - That is, $y_0 \in f(A_2)$.

  ■

- *Claim:* $f(A_1) \subseteq f(A_2) \Longrightarrow A_1 \subseteq A_2$ is not in general true.
  Proof:
  - To prove this we merely have to give one counterexample.
  - Take $X = \{1, 2\}$ and $Y = \{3\}$.
  - Then there can be only *one* map $f : X \to Y$, namely, $f(x) = 3$.
  - Define $A_1 := \{1\}$ and $A_2 := \{2\}$. Then clearly $f(A_1) = f(A_2) = \{3\}$, yet, $A_1 \nsubseteq A_2$.

  ■

- Here is an example of why inverse images are easier to "work with" than images:
  Claim: $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$.
  Proof:

  - $A_1 \cap A_2 \subseteq A_1$, so that $f(A_1 \cap A_2) \subseteq f(A_1)$.
  - $A_1 \cap A_2 \subseteq A_2$, so that $f(A_1 \cap A_2) \subseteq f(A_2)$.
  - Given these two conclusion, we must have $f(A_1 \cap A_2) \subseteq f(A_1 \cap A_2)$.

■

- *Claim:* $f(A_1) \cap f(A_2) \subseteq f(A_1 \cap A_2)$ is *not in general true.*
  Proof:
  - To prove this we merely have to give one counterexample.
  - Take $X = \{1, 2\}$ and $Y = \{3\}$.
  - Then there can be only *one* map $f : X \to Y$, namely, $f(x) = 3$.
  - Define $A_1 := \{1\}$ and $A_2 := \{2\}$.
  - Then $f(A_1) \cap f(A_2) = \{3\}$ yet $A_1 \cap A_2 = \varnothing$ and so $f(\varnothing) = \varnothing$.
  - It is clear that $\{3\} \nsubseteq \varnothing$.

■

1.2 Function Composition

1.2.1 Definition

- Given two functions $f : X \to Y$ and $g : Y \to Z$, we may define a new function $g \circ f : X \to Z$ by the following rule: $x \mapsto g(f(x))$.
- To verify that this definition makes sense, we must make sure that the function is well-defined, that is, that if $x_1 = x_2$ then $(g \circ f)(x_1) = (g \circ f)(x_2)$. This is of course true because f and g are separately well defined (as you can well verify).

1.2.2 Example with Surjectivity

- *Claim:* Let $f_1 : X_1 \to X_2$, $f_2 : X_2 \to X_3$, $\dots$, $f_n : X_n \to X_{n+1}$. If f_i is surjective $\forall i \in \{1, 2, \dots, n\}$ then $f_n \circ f_{n-1} \circ \dots \circ f_1 : X_1 \to X_{n+1}$ is surjective.
  Proof:
  - Let $\alpha \in X_{n+1}$ be given.
  - Our goal is to show that $\exists x \in X_1$ such that $(f_n \circ f_{n-1} \circ \dots \circ f_1)(x) = \alpha$.
  - Because f_n is surjective, $\exists \alpha_n \in X_n$ such that $f_n(\alpha_n) = \alpha$.
  - Because f_{n-1} is surjective, $\exists \alpha_{n-1} \in X_{n-1}$ such that $f_{n-1}(\alpha_{n-1}) = \alpha_n$. That means that $\exists \alpha_{n-1} \in X_{n-1}$ such that $f_n(f_{n-1}(\alpha_{n-1})) = \alpha$.
  - We can continue in this fashion until we reach that there must $\exists \alpha_1 \in X_1$ such that $f_1(\alpha_1) = \alpha_2$, that is, $\exists \alpha_1 \in X_1$ such that $f_n(f_{n-1}(\dots(f_2(f_1(\alpha_1))))) = \alpha$.

■

1.2.3 Example with Injectivity

- *Claim:* Given two functions $f : X \to Y$ and $g : Y \to Z$, if $g \circ f$ is injective, g may in general, be not-injective (even though f is always injective in that scenario).
  Proof:
  - First we show that f is injective:
    * Assume that for given $(x_1, x_2) \in X^2$ we have $f(x_1) = f(x_2)$.
    * To prove the assertion that f is injective we must show that $x_1 = x_2$.
    * Because $f(x_1) = f(x_2)$, then necessarily as g is well defined we have $g(f(x_1)) = g(f(x_2))$, that is, $(g \circ f)(x_1) = (g \circ f)(x_2)$.
    * But $g \circ f$ is injective, which means that $x_1 = x_2$.
  - Suffice to give one counterexample to prove this claim:
    * Let $X = [0, \infty)$, $Y = \mathbb{R}$ and $Z = \mathbb{R}$.
    * Define $f : X \to Y$ by $x \mapsto x$ and $g : Y \to Z$ by $y \to y^2$.
    * Then clearly f is injective (as it is just the identity map) and as we've seen before g is *not* injective.
    * However, $g \circ f$ is injective as you can well verify explicitly:
      · Assume for two x_1 and x_2 in X ($(x_1, x_2) \in X^2$) we have that $(g \circ f)(x_1) = (g \circ f)(x_2)$.
      · That means that $g(f(x_1)) = g(f(x_2))$.
      · Because f is just the identity map, that means $g(x_1) = g(x_2)$, or ${x_1}^2 = {x_2}^2$.
      · But x_1 and x_2 are positive numbers as we've defined X, so that $x_1 = x_2$.
    * Conclusion: Even though g is not injective, $g|_{f(X)} : f(X) \to Z$ (the restriction of g to the image of f) is going to be injective, and this is the intuitive piece we were missing in this theorem.

1.3 Cardinalities

1.3.1 The Set of Finite Subsets of $\mathbb{R}$ Has the Same Cardinality as $\mathbb{R}$

- Define $S := \{ A \subset \mathbb{R} \mid |A| \in \mathbb{N} \}$
- *Claim*: $|\mathbb{R}| = |S|$
  Proof:
  - Define $S_n := \{ A \subset \mathbb{R} \mid |A| = n \}$ for all $n \in \mathbb{N}$.
  - Then clearly, $S = \bigcup_{n \in \mathbb{N}} S_n$.
  - *Claim:* $|\mathbb{R}| \leq |S|$.
    Proof:
    * $S_1 = \{ A \subset \mathbb{R} \mid |A| = 1 \} = \{ \{r\} \subset \mathbb{R} \mid r \in \mathbb{R} \}$ so that $|S_1| = |\mathbb{R}|$ and thus as $S_1 \subseteq S$, $|S| \geq |\mathbb{R}|$.
  - *Claim:* $|\mathbb{R}| \geq |S|$.
    Proof:
    * Let $A \in S_n$ be given. That is, A is a set of real numbers with n elements.
    * Then we can write these elements as $a_1, a_2, \dots, a_n$ and WLOG can assume that $a_1 < a_2 < \dots < a_n$.
    * Define a map $f : S_n \to \mathbb{R}^n$ by $A \mapsto (a_1, a_2, \dots, a_n)$. (That is, by the order of the elements).

* *Claim:* f is well defined.
  Proof:
    - Clearly by ordering the elements we remove any ambiguity about the representation of the sets: if two sets are equal they will necessarily lead to the same n-tuplet in $\mathbb{R}^n$.
* *Claim:* f is injective.
  Proof:
    - By the same argumentation as above it is clear that if two n-tuplets are equal in the image of f they must have come from the same finite subset of $\mathbb{R}$.
* Because f is injective, we necessarily have $|S_n| \leq |\mathbb{R}^n|$.
* *Claim:* $|\mathbb{R}^n| = |\mathbb{R}|$.
    - *Claim:* $|(0,1)| = |\mathbb{R}|$. *Proof:* the function $\tan\left(\pi x - \frac{\pi}{2}\right)$ from $(0,\,1)$ to $\mathbb{R}$ is a bijection. (Proof as homework.)
    - *Claim:* $|(0,\,1)| = \left|(0,\,1)^2\right|$.
      Proof:
      1. For every real number between 0 and 1, pick one particular convention for a decimal representation. (For example, pick 0.01 always instead of $0.009999999999\ldots$)
      2. Define a map $g : (0,\,1)^2 \to (0,\,1)$ by the relation $(0.a1a2a3\ldots,\; 0.b1b2b3\ldots) \mapsto 0.a1b1a2b2a3b3\ldots$.
      3. *Claim:* g is injective. *Proof:* If $g\left((0.a1a2a3\ldots,\; 0.b1b2b3\ldots)\right) = g\left((0.c1c2c3\ldots,\; 0.d1d2d3\ldots)\right)$ then $0.a1b1a2b2a3b3\cdots =$ $0.c1d1c2d2c3d3\ldots$, in which case it is clear that $a_i = c_i$ and $b_i = d_i$ for all i, and so really the two pairs are equal.
      4. *Claim:* g is surjective. *Proof:* Given a random number $0.x1x2x3\cdots \in (0,\,1)$, Take the pair $(0.x1x3x5\ldots,\; 0.x2x4x6\ldots)$ which will map to it.
    - Returning to the actual proof that $|\mathbb{R}^n| = |\mathbb{R}|$, we perform induction on n.
    - The induction basis is $n = 2$. For this case, we have that

$$\begin{aligned} \left|\mathbb{R}^2\right| &= |\mathbb{R} \times \mathbb{R}| \\ &= |(0,\,1) \times (0,\,1)| \\ &= |(0,\,1)| \\ &= |\mathbb{R}| \end{aligned}$$

    - Assume that $|\mathbb{R}^n| = |\mathbb{R}|$ for *some* $n \in \{2,\,3,\,\ldots\}$.
    - Then

$$\begin{aligned} \left|\mathbb{R}^{n+1}\right| &= |\mathbb{R}^n \times \mathbb{R}| \\ &= |\mathbb{R} \times \mathbb{R}| \\ &= |\mathbb{R}| \end{aligned}$$

    - Thus by the principle of induction we have that $|\mathbb{R}^n| = |\mathbb{R}|$ for all $n \in \{2,\,3,\,\ldots\}$. Of course the statement is true for $n = 1$!
* Thus we have $|S_n| \leq |\mathbb{R}|$.
* But $|S| = \left|\bigcup_{n\in\mathbb{N}} S_n\right|$, and all the S_n's are disjoint. Because each has $|S_n| \leq |\mathbb{R}|$, we may estimate the cardinality of $\left|\bigcup_{n\in\mathbb{N}} S_n\right| \overset{*}{\leq} \left|\bigcup_{n\in\mathbb{N}} \{n\} \times \mathbb{R}\right| = |\mathbb{N} \times \mathbb{R}| \leq |\mathbb{R} \times \mathbb{R}| = |\mathbb{R}|$, where in $*$ we used the fact that all the S_n's are disjoint, so we can say the cardinality of the union is necessarily smaller than the cardinality of $|\mathbb{N}|$ disjoint copies of $\mathbb{R}$.

– Because $|S| \leq |\mathbb{R}|$ and $|S| \geq |\mathbb{R}|$, using the Schroeder-Bernstein theorem we conclude that $|S| = |\mathbb{R}|$.

1.3.2 $|\mathbb{N}|$ is the Smallest Possible Infinite Cardinality

- *Claim:* For any infinite set X, $\exists$ an injection $\mathbb{N}\to X$ (which shows that $|\mathbb{N}|\leq|X|$.
  Proof:
  - Because X is infinite, it is not empty, so $\exists x_0\in X$. Define $f(0):=x_0$.
  - Because X is infinite, $X\backslash\{x_0\}$ is not empty, so $\exists x_1\in X$. Define $f(1):=x_1$.
  - Because X is infinite, $X\backslash\{x_0,\,x_1\}$ is not empty, so $\exists x_2\in X$. Define $f(2):=x_2$.
  - ...
  - Because X is infinite, $X\backslash\{x_0,\,x_1,\,x_2,\,\dots,\,x_n\}$ is not empty, so $\exists x_{n+1}\in X$. Define $f(n+1):=x_{n+1}$. This is possible due to the axiom of dependent choice which says that you can make countably many choices, each depending on the previous ones you made. In fact we can prove that DC stems from AC:
  - *Claim: AC* implies *DC* (axiom of choice implies dependent axiom of choice)
    Proof:
    * Define for each $A\subset X$, $R(A):=\{\,B\subset X\mid A\subset B\wedge|B|=|A|+1\,\}$.
    * By assumption we assume $R(A)\neq\varnothing\forall A\subset X$. This amounts to the information that X is infinite.
    * Using the axiom of choice, $\exists f:\mathcal{P}(X)\to\mathcal{P}(X)$ such that $f(A)\in R(A)\,\forall A\subset X$.
    * So for any $A\subset X$, the sequence $(A_n)_{n\in\mathbb{N}}:=f^n(A)$, where f^n denotes composition of f n times on itself, has $A_n\in R(A_{n-1})$ for all $n\in\mathbb{N}$.
  - Clearly f is injective because we always pick our next value from the complement of the previous values.

1.4 Specific Tips

1.4.1 Question 4

- Typo in question: You will get an injective map from $\mathbb{N}\to X$ which shows that any infinite set has $|\mathbb{N}|\leq|X|$, that is, $|\mathbb{N}|$ is the smallest cardinality. Point of the exercise is (as far as I can see) to repeat the proof above but without the full axiom of choice, rather, instead, just the *countable* axiom of choice.

- (a): Use induction to prove the statement for all $n\in\mathbb{N}$. Use the fact that X is infinite, so that in particular it is not empty and has as many elements as you want, even after you remove a finite number from it.

- (b): Use the axiom of choice (the infinite product of nonempty sets is nonempty) to ascertain that there is such a map, because the product $\prod_{k\in\mathbb{N}}\{\,A\subset X\mid|A|=2^k\,\}$ is non-empty (because the sets in the product, $\{\,A\subset X\mid|A|=2^k\,\}$, are not empty using (a)). But now we have:

$$\begin{aligned}\prod_{k\in\mathbb{N}}\{\,A\subset X\mid|A|=2^k\,\} &\equiv \{\,A\subset X\mid|A|=2^0\,\}\times\{\,A\subset X\mid|A|=2^1\,\}\times\{\,A\subset X\mid|A|=2^2\,\}\times\dots\\ &= \left\{\,(A,\,B,\,C,\,\dots)\in\mathcal{P}(X)^{\mathbb{N}}\,\middle|\,|A|=1,\,|B|=2,\,|C|=4,\,\dots\right\}\end{aligned}$$

  We can identify any infinite product $(a,\,b,\,c,\,\dots)\in X^{\mathbb{N}}$ as a map $A:\mathbb{N}\to X$ (don't confuse A as a set or Aas a map–sorry for the duplicity). A map is just a rule that associates with every number $n\in\mathbb{N}$ an element of our set X, and this is exactly what an infinite product does as well: given a number $n\in\mathbb{N}$, we have associated with it an element of X: for $n=0$ we have associated a, for $n=1$ we have associated b and so on. Thus

$$\left\{\,(A,\,B,\,C,\,\dots)\in\mathcal{P}(X)^{\mathbb{N}}\,\middle|\,|A|=1,\,|B|=2,\,|C|=4,\,\dots\right\} = \{\,A:\mathbb{N}\to\mathcal{P}(X)\mid|A(k)|=2^k\forall k\in\mathbb{N}\,\}$$

- (c): Define $S(k) := A(k) \setminus \left(\bigcup_{j=1}^{k-1} A(k)\right)$. Make estimates on the size of $\bigcup_{j=1}^{k-1} A(k)$ to show that $S(k)$ is nonempty. Then using the axiom of choice we know that $\prod_{k\in\mathbb{N}} S(k)$ is not empty, and an element in this product is just a function $f : \mathbb{N} \to X$. (think why and use the same reasoning as above).

1.4.2 Question 5

- (a): Show that the I_n's build nested intervals, and as seen on the lecture, nested intervals correspond to unique real numbers. As a result, g is (well) defined exactly as that one unique number which is in the intersection of all nested intervals.

- (b): Pick a random real number $x \in [0, 1]$. Show that we can define a binary sequence that will give rise to nested intervals that "close in" on this number. This will show surjectivity. To define the sequence, think about how these intervals "close in" on a number: for example, for the first number in the sequence pick $a_0 := 0$. For the second, a_1, if $x \in (0, \frac{1}{2}]$ then pick $a_1 := 0$ and otherwise pick $a_1 := 1$. How to generalize this? Is this choice always canonical? (this will answer injectivity)

1.4.3 Question 6

- (c) and (d): Use this "trick":

$$\begin{aligned}
\frac{a+ib}{c+id} &= \left(\frac{a+ib}{c+id}\right)\underbrace{\left(\frac{c-id}{c-id}\right)}_{1} \\
&= \frac{(a+ib)(c-id)}{c^2-d^2} \\
&= \frac{ac+bd+i(bc-ad)}{c^2-d^2} \\
&= \frac{ac+bd}{c^2-d^2} + i\frac{bc-ad}{c^2-d^2}
\end{aligned}$$

- (e) and (f): Plug in $z = a + ib$ where $(a, b) \in \mathbb{R}^2$ and find a and b. Make the necessary computations to get a complex equation of the form: $a + ib = c + id$. This will give you then *two* real equations for two real variables, a and b (which correspond to one complex variable, z).

1.4.4 Question 7

- Translate the conditions into conditions on two real variables a and b, and then sketch the solutions on $\mathbb{R}^2$.

- (e): What is the imaginary part of the absolute value of a complex number? What kind of values can the absolute value take anyway?

2 Homework Sheet Number 1

- Common mistakes:
  - Review induction!! In particular, what's "$n = n + 1$"? Make sure you understand: (a) How to formulate a proof by induction and (b) Why is the formulation the way it is (stems from a particular logic).
  - Set builder notation: distinction between *sets* and *formulas*.

- How to prove that two sets are equal (especially *not* using Venn diagrams).

1. Question 1

   (a) Define a statement $A(n)$ for every $n \in \mathbb{N}$, such that $A(n)$ is true iff $\sum_{k=1}^{n} k^2 = \frac{1}{6}[n(n+1)(2n+1)]$.
   Claim: $A(n)\ \forall n \in \mathbb{N}$
   Proof:

      i. *Claim:* $A(1)$.
      Proof:
         - $A(1)$ says that $\sum_{k=1}^{1} k^2 = \frac{1}{6}[1(1+1)(2 \cdot 1+1)]$ which is of course true.

      ii. *Claim:* $[A(n) \Longrightarrow A(n+1)]\ \forall n \in \mathbb{N}$.
      Proof:
         - Let $n \in \mathbb{N}$ be given, and assume that $A(n)$.
         - $A(n+1)$ says that $\sum_{k=1}^{n+1} k^2 = \frac{1}{6}[(n+1)(n+2)(2(n+1)+1)]$.
         - Start from the left hand side of the equation:

$$\begin{aligned}
\sum_{k=1}^{n+1} k^2 &= \sum_{k=1}^{n} k^2 + (n+1)^2 \\
&= \frac{1}{6}[n(n+1)(2n+1)] + (n+1)^2 \\
&= (n+1)\left\{\frac{1}{6}[n(2n+1)] + (n+1)\right\} \\
&= \frac{1}{6}(n+1)[n(2n+1) + 6(n+1)] \\
&= \frac{1}{6}\left[(n+1)\left(2n^2+7n+6\right)\right] \\
&= \frac{1}{6}[(n+1)(n+2)(2n+3)] \\
&= \frac{1}{6}[(n+1)(n+2)(2(n+1)+1)]
\end{aligned}$$

         - Thus we have reached the right hand side of the desired equation.

   ■

2.

3.

4.

5. Question 5

   (a)

   (b) $\{\, x \in \mathbb{R} \mid |x-2| \geq |x| - 2 \,\} = ?$
      - The inequality: $|x-2| \geq |x| - 2$ is equivalent to $|x-2| + 2 \geq |x|$.
      - Case 1: $x \geq 2$, then we have $x - 2 + 2 \geq x$ which is always true.
      - Case 2: $0 \leq x \leq 2$, then we have $-x + 2 + 2 \geq x$ which means $2x \leq 4$ or $x \leq 2$ which we anyway have.

- Case 3: $x < 0$, then we have $-x + 2 + 2 \geq -x$ which is always true.
- General conclusion from this exercise: whenever you are a bit confused about many absolute values, just separate into all possible cases.

(c)

(d) $1 \leq |x| + |y| \leq 2$.

- Compare with $1 \leq x^2 + y^2 \leq 2$.
- In general, separate into four different cases:

i. $x \geq 0$ and $y \geq 0$, then we have $1 \leq x + y \leq 2$, which means $1 - x \leq y \leq 2 - x$.

ii. Like that all four possible cases.

6.

7. Question 7

- Let L be the set of all books. Thus $|L|$ is the number of all books in the library.
- $\forall B \in L$, $|B|$ is the number of words in a book B.
- *Claim:* There is at least one empty book.
  Proof:
  - Assume otherwise. That is, there is *no* empty book.
  - We know that the number of books, $|L|$, is larger than the sum of words in all books, $\sum_{B \in L} |B|$. That is, $|L| > \sum_{B \in L} |B|$.
  - If no book is empty, then $|B| \geq 1$ for all $B \in L$. Then $\sum_{B \in L} |B| > \sum_{B \in L} 1 = |L|$.
  - Thus we find that $|L| > |L|$, a contradiction.
  - Thus, there must be at least one empty book in the library.
- *Claim:* There must be less than three books in the library.
  Proof:
  - We know that there are no two books in the library with the same number of words.
  - Thus, order all the books by word number, say, 0, 3, 50, 3000, 3500, ...
  - The only thing we know is that
    * this sequence starts with zero,
    * that there are $|L|$ numbers in it, and
    * that each item in the sequence must be bigger by at least one than its predecessor.
  - Using these constraints we can think of the "worst case" scenario, meaning, the case in which there is the minimal number of words in every book: 0, 1, 2, ... Any actual sequence of number of words of books in the library will be bigger than that.
  - Thus, $\sum_{B \in L} |B| \geq \sum_{k=0}^{|L|-1} k = \frac{1}{2}(|L| - 1)|L|$.
  - But we know that $|L| > \sum_{B \in L} |B|$ so that $|L| > \frac{1}{2}(|L| - 1)|L|$.
  - This equation may only be fulfilled when $0 < |L| < 3$.
  - Thus we can only have either one or two books in the library.
  - But one of the two books has to be empty, so if there are two books, the next one must have exactly one word to maintain the equation: $\underbrace{|L|}_{2} > \underbrace{|B_1|}_{0} + |B_2|$.