26 Appendix: "Proof machine"

This section is the salvation for a student of mathematics.

26.1 Andante cantabile

26.1.1 Memories

It was in the second semester of my undergraduate education at MIT that I first met pure mathematics, open and closed sets, the book "Baby Rudin", and Warren Ambrose. The course was '18.100 Mathematical Analysis'. Warren Ambrose had a great effect on me. Somehow we had a one-to-one conversation where we both confessed that our true love was music and that we were doing math only as a backup. At the time, I was still far from being a professional mathematician and he was a famous geometer nearing the end of his career and his life (it was 1984 and he died in 1995 at the age of 81). He told me that he had been a jazz trumpet player but an accident had made him unable to play properly and so he had pursued mathematics for a profession. His exams (two midterm exams and a final) were all 24 hour open-book closed-friend take-home tests: 10 questions, true or false, graded 1 if correct, -1 if incorrect, and 0 if not answered. The average score across the class (about 20 students) was often around 0. But this mechanism taught you better than any other what proof meant – if you were unable to provide a proof you believed in then you risked getting -1 for that question. The questions were always very interesting. I carried those questions around for years until sometime in 2012 when I accidentally left them in a classroom and, when I came back to find them an hour later, they were gone.

26.1.2 Assume the Ifs and To show the Thens: "Proof machine"

The first courses I had that required me to start constructing proofs (Mathematical Analysis, Abstract algebra, Topology) were tough for me. I couldn't figure out the magic trick that made some people able to do this. By the time I started graduate school I still hadn't figured out this magic and I thought it likely that without it it would be impossible for me to succeed in obtaining a PhD in mathematics. On the other hand I began to notice that, in combinatorics particularly, if I knew that I could make some bijection or other then I was *absolutely sure* that I could make it and there was something more than just wishy-washy hand waving that I was doing to have this certainty. I was just starting to get the hang of it.

It was when I was a postdoc that I realized that most of mathematics is just mechanical work, and the bright ideas that are needed are few and far between. This gave me confidence as I was sure that I had the diligence and endurance to do mechanical work, and I was also pretty certain that if any actual "talent" was going to be required then I wasn't going to be a successful mathematician.

Just at that moment I got assigned to teach the undergraduate Abstract Algebra course (at Univ. of Wisconsin–Madison) and so I needed to figure out how to explain to my students how they too could do the necessary proofs. That was the catalyst for me to formulate the mechanism that I now call "proof machine".

As I have progressed in a career as a professional research mathematician I have been amazed to observe how many times "proof machine" has saved me, provided the direction, guided me to where I might have to think, clarified where I didn't need to waste effort thinking, provided the proof and protected me from making mistakes.

"Proof machine" was also the key that unlocked the mysterious world of writing and changed me from a teenager who hated English class, any kind of writing and especially term papers, into a versatile writer (at least in the cases when I do the writing carefully and thoroughly and with the same structural framework that I use when I do a proof in "proof machine" in mathematics). I am always struck by how helpful "proof machine" is for getting out good writing (letters, reports, reviews, papers, memos, emails, etc).

I am continually amazed at how useful "proof machine" is in my daily life and meetings, in helping me be organised and efficient, helping me to get to the core of the issue as necessary, and helping me to optimize impact and productivity for effort expended. "Proof machine" is a skill (not a talent) which is learned by practice (and more practice and more practice) in the same way that one develops skill and facility on a musical instrument by lots of practice.

My hope is that I can teach"proof machine" to as many of my students as I can so that they can also benefit from this wonderful tool in their lives and careers. After all, it is really easy: To prove "If A then B", *Assume* the ifs and *To show* the thens, and that's about all there is to it. The rest is just practice.

26.2 The grammar of mathematics

- **Definitions** are the foundation of mathematics.
- **Theorems** are the landmarks of mathematics.
- **Proofs** are the explanation of mathematics.

Learning to read, write and speak mathematics is a skill that anyone can learn. Like all languages, it requires lots of practice to use it fluently.

Like all languages, the grammar of quality mathematical communication is very rigid.

It is **impossible** to prove a statement without being able to write down the definitions of all the terms in the statement.

The grammar of a definition is:

A noun is a _______ such that

(a) If _______ then _______, and

(b) If _______ then _______, and

(c) If _______ then _______, and ...

Let $\mathbb{F}$ be field and let V and W be $\mathbb{F}$-vector spaces.
A linear transformation from V to W is a function $f\colon V \to W$ such that

(a) If $v_1, v_2 \in V$ then $f(v_1 + v_2) = f(v_1) + f(v_2)$,

(b) If $c \in \mathbb{F}$ and $v \in V$ then $f(cv) = cf(v)$.

An adjective is most conveniently defined by putting it in the form of a noun:

A adjective noun is a noun such that

(a) If _______ then _______, and

(b) If _______ then _______, and

(c) If _______ then _______, and ...

An injective function is a function $f\colon S \to T$ such that

(a) If $\underline{s_1, s_2 \in S \text{ and } s_1 \neq s_2}$ then $\underline{f(s_1) \neq f(s_2)}$.

Sometimes definitions of adjectives take the form:

Let S be a noun.
A noun S is adjective if S satisfies

(a) If _______ then _______, and

(b) If _______ then _______, and

(c) If _______ then _______, and ...

Let $f\colon S \to T$ be a function.
A function $f\colon S \to T$ is injective if f satisfies

(a) If $\underline{s_1, s_2 \in S \text{ and } s_1 \neq s_2}$ then $\underline{f(s_1) \neq f(s_2)}$.

The words "let" and "assume" are synonyms for "if". The grammar of a lemma, proposition or theorem (or any other statement) is:

If _______ then _______.

Two special constructions in mathematical language are:

There exists _______ such that _______.

and

There exists a unique _______ such that _______.

26.3 How to do Proofs: “Proof Machine”

There *is* a certain “formula” or method to doing proofs. Some of the guidelines are given below. The most important factor in learning to do proofs is practice, just as when one is learning a new language.

1. There are very few words needed in the structure of a proof. Organized in rows by synonyms they are:

   ```
   To show
   Assume, Let, Suppose, Define, If
   Since, Because, By
   Then, Thus, So
   There exists, There is
   Recall, We know, But
   ```

   Do not use ‘`for all`’ or ‘`for each`’. These can always be replaced by ‘`if`’ to achieve greater clarity, accuracy and efficiency.

   Do not use the phrase ‘`for some`’. It can always be replaced by ‘`There exists`’ to achieve greater clarity, accuracy and efficiency.

2. The overall structure of a proof is a block structure like an outline. For example:

   ```
   To show:  If __A__ then __B__ and __C__.
       Assume:  __A__.
   [itemsep=-0.2em]
   To show:  (a) __B__.
             (b) __C__.
   (a) To show:  __B__.
           _______.
           _______.
           _______.
       Thus __B__.
   (b) To show:  __C__.
           _______.
           _______.
           _______.
       Thus __C__.
   So __B__ and __C__.
   So, if __A__ then __B__ and __C__.
   ```

3. Any proof or section of proof begins with one of the following:

   ```
   (a) To show:  If __A__ then __B__.
   (b) To show:  There exists __C__ such that __D__.
   (c) To show:  __E__.
   ```

4. Immediately following this, the next step is

Case (a) Assume the ifs and 'To show' the thens. The next lines are

- ∘ Assume __A__.
- ∘ To show: __B__.

Case (b) To show an object exists you must find it. The next lines are

- ∘ Define __xxx = ________.
- ∘ To show: __xxx__ satisfies __D__.

Case (c) Rewrite the statement in __E__ by using a definition. The next line is

- ∘ To show: __E'__.

There are some kinds of proofs which have a special structure.

(E) Proofs of equality: LHS=RHS.

To show: __A=B__.

Left Hand Side: __A= $\cdots$____
= $\cdots$
= $\cdots$
= $\cdots$
= expression____

Right Hand Side: __B= $\cdots$____
= $\cdots$
= $\cdots$
= $\cdots$
= THE SAME expression____

(F) Counterexamples: Proofs of falseness

To show that a statement, "If ____ then ____", is false you *must* give an example.

To show: There exists a __xxx__ such that
(a) xxx satisfies the ifs of the statement that you are showing is false,
(a) xxx satisfies the opposite of some assertion in the thens of the statement that you are showing is false.

(U) Proofs of uniqueness.

To show that an object is unique you must show that if there are two of them then they are really the same.

To show: A THING is unique.
Assume X_1 and X_2 are both THINGs.
To show: $X_1 = X_2$.

(I) Proofs by induction.

A statement to be proved by induction *must* have the form

```
If $n$ is a positive integer then __A__.
```

The proof by induction should have the form

```
Proof by induction.
Base case:
To show:  If $n = 1$ then __A__.
       ______.
       ______.
       ______.
Thus, if $n = 1$ then __A__.
Induction step:
Let $\ell$ be a positive integer and assume that if $n$ is a positive integer and $n <$
$\ell$ then __A__.
To show: __A__.
```

The mechanics of proof by induction is an unwinding of the *definition* of $\mathbb{Z}_{>0}$.

(CP) Proofs by contrapositive.

```
To show:  If __A__ then __B__.
To show:  If __not B__ then __not A__.
```

(BAD) Proofs by contradiction.

(*) Assume the opposite of what you want to show.
______.
______.
______.
End up showing the opposite of some assumption (not necessarily the (*) assumption).
Contradiction to specify exactly what assumption is being contradicted.
Thus assumption (*) is wrong and what you want to show is true.

PROOFS BY CONTRADICTION ARE STRONGLY DISCOURAGED. In all known cases they can be replaced by a proof by contrapositive for greater clarity, direction and efficiency.

26.4 Example proofs

The following example proofs have been chosen because they are results that are often assumed, are needed for many topics in algebra and analysis and topology and are rarely proved carefully in an undergraduate curriculum; facts like, if $a \neq 0$ then $a^2 > 0$. These often seem "obvious", until you meet that first example, like a field witih 5 elements, where $2 \neq 0$ and $2^2 = -1$. After getting over the initial shock, then one begins to wonder why such a fact might ever be true, and how it might be proved when it is. It is proved in Proposition 26.4(b), below.

26.4.1 An inverse function to f exists if and only if f is bijective.

Theorem 26.1. *Let $f\colon S \to T$ be a function. The inverse function to f exists if and only if f is bijective.*

Proof.

$\Rightarrow$: Assume $f\colon S \to T$ has an inverse function $f^{-1}\colon T \to S$.
To show: (a) f is injective.
(b) f is surjective.

(a) Assume $s_1, s_2 \in S$ and $f(s_1) = f(s_2)$.
To show: $s_1 = s_2$.

$$s_1 = f^{-1}f(s_1)) = f^{-1}f(s_2)) = s_2.$$

So f is injective.

(b) Let $t \in T$.
To show: There exists $s \in S$ such that $f(s) = t$.
Let $s = f^{-1}(t)$.
Then

$$f(s) = f(f^{-1}(t)) = t.$$

So f is surjective.

So f is bijective.

$\Leftarrow$: Assume $f\colon S \to T$ is bijective.
To show: f has an inverse function.
We need to define a function $\varphi\colon T \to S$.
Let $t \in T$.
Since f is surjective there eists $s \in S$ such that $f(s) = t$.
Define $\varphi(t) = s$.
To show: (a) φ is well defined.
(b) φ is an inverse function to f.

(a) To show: (aa) If $t \in T$ then $\varphi(t) \in S$.
(ab) If $t_1, t_2 \in T$ and $t_1 = t_2$ then $\varphi(t_1) = \varphi(t_2)$.

(aa) This follows from the definition of φ.
(ab) Assume $t_1, t_2 \in T$ and $t_1 = t_2$.
Let $s_1, s_2 \in S$ such that $f(s_1) = t_1$ and $f(s_2) = t_2$.

Since $t_1 = t_2$ then $f(s_1) = f(s_2)$.
Since f is injective this implies that $s_1 = s_2$.
So $\varphi(t_1) = s_1 = s_2 = \varphi(t_2)$.

So φ is well defined.

(b) To show: (ba) If $s \in S$ then $\varphi(f(s)) = s$.
(bb) If $t \in T$ then $f(\varphi(t)) = t$.

(ba) This follows from the definition of φ.

(bb) Assume $t \in T$.
Let $s \in S$ be such that $f(s) = t$.
Then

$$f(\varphi(t)) = f(s) = t.$$

So $\varphi \circ f$ and $f \circ \varphi$ are the identity functions on S and T, respectively.

So φ is an inverse function to f.

□

26.4.2 An equivalence relation on S and a partition of S are the same data.

Let S be a set.

- A *relation* $\sim$ *on* S is a subset $R_\sim$ of $S \times S$. Write $s_1 \sim s_2$ if the pair (s_1, s_2) is in the subset $R_\sim$ so that
$$R_\sim = \{(s_1, s_2) \in S \times S \mid s_1 \sim s_2\}.$$
- An *equivalence relation* on S is a relation $\sim$ on S such that
  (a) if $s \in S$ then $s \sim s$,
  (b) if $s_1, s_2 \in S$ and $s_1 \sim s_2$ then $s_2 \sim s_1$,
  (c) if $s_1, s_2, s_3 \in S$ and $s_1 \sim s_2$ and $s_2 \sim s_3$ then $s_1 \sim s_3$.

Let $\sim$ be an equivalence relation on a set S and let $s \in S$. The *equivalence class of* s is the set

$$[s] = \{t \in S \mid t \sim s\}.$$

A *partition of a set* S is a collection $\mathcal{P}$ of subsets of S such that

(a) If $s \in S$ then there exists $P \in \mathcal{P}$ such that $s \in P$, and
(b) If $P_1, P_2 \in \mathcal{P}$ and $P_1 \cap P_2 \neq \emptyset$ then $P_1 = P_2$.

Theorem 26.2.

(a) If S is a set and let $\sim$ be an equivalence relation on S then

$$\text{the set of equivalence classes of} \sim \qquad \text{is a partition of } S.$$

(b) If S is a set and $\mathcal{P}$ is a partition of S then

$$\text{the relation defined by} \qquad s \sim t \quad \text{if } s \text{ and } t \text{ are in the same } P \in \mathcal{P}$$

is an equivalence relation on S.

Proof.

(a) To show: (aa) If $s \in S$ then s is in some equivalence class.

(ab) If $[s]\cap[t]\ne\emptyset$ then $[s]=[t]$.

(aa) Let $s\in S$.
Since $s\sim s$ then $s\in[s]$.

(ab) Assume $[s]\cap[t]\ne\emptyset$.
To show: $[s]=[t]$.
Since $[s]\cap[t]\ne\emptyset$ then there is an $r\in[s]\cap[t]$.
So $s\sim r$ and $r\sim t$.
By transitivity, $s\sim t$.
To show: (aba) $[s]\subseteq[t]$.
(abb) $[t]\subseteq[s]$.

(aba) Assume $u\in[s]$.
Then $u\sim s$.
We know $s\sim t$.
So, by transitivity, $u\sim t$.
Therefore $u\in[t]$.

So $[s]\subseteq[t]$.

(aba) Assume $v\in[t]$.
Then $v\sim t$.
We know $t\sim s$.
So, by transitivity, $v\sim s$.
Therefore $v\in[s]$.

So $[t]\subseteq[s]$.
So $[s]=[t]$.

So the equivalence classes partition S.

(b) To show: $\sim$ is an equivalence relation, i.e. that $\sim$ is reflexive, symmetric and transitive.
To show: (ba) If $s\in S$ then $s\sim s$.
(bb) If $s\sim t$ then $t\sim s$.
(bc) If $s\sim t$ and $t\sim u$ then $s\sim u$.

(ba) Since s and s are in the same S_α then $s\sim s$.

(bb) Assume $s\sim t$.

Then s and t are in the same S_α.

So $t\sim s$.

(bb) Assume $s\sim t$ and $t\sim u$.

Then s and t are in the same S_α and t and u are in the same S_α.

So $s\sim u$.

So $\sim$ is an equivalence relation.

□

26.4.3 Identities in a field

A *field* is a set $\mathbb{F}$ with functions

$$\begin{array}{ccc}\mathbb{F}\times\mathbb{F} & \longrightarrow & \mathbb{F}\\ (a,b) & \longmapsto & a+b\end{array}\qquad\text{and}\qquad\begin{array}{ccc}\mathbb{F}\times\mathbb{F} & \longrightarrow & \mathbb{F}\\ (a,b) & \longmapsto & ab\end{array}$$

such that

(Fa) If $a, b, c \in \mathbb{F}$ then $(a+b)+c = a+(b+c)$,

(Fb) If $a, b \in \mathbb{F}$ then $a+b = b+a$,

(Fc) There exists $0 \in \mathbb{F}$ such that

$$\text{if } a \in \mathbb{F} \quad \text{then} \quad 0+a = a \text{ and } a+0 = a,$$

(Fd) If $a \in \mathbb{F}$ then there exists $-a \in \mathbb{F}$ such that $a+(-a) = 0$ and $(-a)+a = 0$,

(Fe) If $a, b, c \in \mathbb{F}$ then $(ab)c = a(bc)$,

(Ff) If $a, b, c \in \mathbb{F}$ then

$$(a+b)c = ac+bc \qquad \text{and} \qquad c(a+b) = ca+cb,$$

(Fg) There exists $1 \in \mathbb{F}$ such that

$$\text{if } a \in \mathbb{F} \quad \text{then} \quad 1 \cdot a = a \text{ and } a \cdot 1 = a,$$

(Fh) If $a \in \mathbb{F}$ and $a \neq 0$ then there exists $a^{-1} \in \mathbb{F}$ such that $aa^{-1} = 1$ and $a^{-1}a = 1$,

(Fi) If $a, b \in \mathbb{F}$ then $ab = ba$.

Proposition 26.3. *Let $\mathbb{F}$ be a field.*

(a) If $a \in \mathbb{F}$ then $a \cdot 0 = 0$.

(b) If $a \in \mathbb{F}$ then $-(-a) = a$.

(c) If $a \in \mathbb{F}$ and $a \neq 0$ then $(a^{-1})^{-1} = a$.

(d) If $a \in \mathbb{F}$ then $a(-1) = -a$.

(e) If $a, b \in \mathbb{F}$ then $(-a)b = -ab$.

(f) If $a, b \in \mathbb{F}$ then $(-a)(-b) = ab$.

Proof.

(a) Assume $a \in \mathbb{F}$.

$$\begin{aligned} a \cdot 0 &= a \cdot (0+0), \quad \text{by (Fc)}, \\ &= a \cdot 0 + a \cdot 0, \quad \text{by (Ff)}. \end{aligned}$$

Add $-a \cdot 0$ to each side and use (Fd) to get $0 = a \cdot 0$.

(b) Assume $a \in \mathbb{F}$.
By (Fd),

$$-(-a)+(-a) = 0 = a+(-a).$$

Add $-a$ to each side and use (Fd) to get $-(-a) = a$.

(c) Assume $a \in \mathbb{F}$ and $a \neq 0$.
By (Fh),

$$(a^{-1})^{-1} \cdot a^{-1} = 1 = a \cdot a^{-1}.$$

Multiply each side by a and use (Fh) and (Fg) to get $(a^{-1})^{-1} = a$.

(d) Assume $a \in \mathbb{F}$.
By (Ff),

$$a(-1) + a \cdot 1 = a(-1+1) = a \cdot 0 = 0,$$

where the last equality follows from part (a).

So, by (Fg), $a(-1)+a=0$.

Add $-a$ to each side and use (Fd) and (Fc) to get $a(-1)=-a$.

(e) Assume $a,b\in\mathbb{F}$.

$$\begin{aligned}(-a)b+ab &= (-a+a)b, \quad \text{by (Ff)},\\ &= 0\cdot b, \quad \text{by (Fd)},\\ &= 0, \quad \text{by part (a)}.\end{aligned}$$

Add $-ab$ to each side and use (Fd) and (Fc) to get $(-a)b=-ab$.

(f) Assume $a,b\in\mathbb{F}$.

$$\begin{aligned}(-a)(-b) &= -(a(-b)), \quad \text{by (e)},\\ &= -(-ab), \quad \text{by (e)},\\ &= ab, \quad \text{by part (b)}.\end{aligned}$$

□

26.4.4 Identities in an ordered field

An *ordered field* is a field $\mathbb{F}$ with a total order $\le$ such that

(OFa) If $a,b,c\in\mathbb{F}$ and $a\le b$ then $a+c\le b+c$,

(OFb) If $a,b\in\mathbb{F}$ and $a\ge 0$ and $b\ge 0$ then $ab\ge 0$.

Proposition 26.4. *Let $\mathbb{F}$ be an ordered field.*

(a) If $a\in\mathbb{F}$ and $a>0$ then $-a<0$.

(b) If $a\in\mathbb{F}$ and $a\ne 0$ then $a^2>0$.

(c) $1\ge 0$.

(d) If $a\in\mathbb{F}$ and $a>0$ then $a^{-1}>0$.

(e) If $a,b\in\mathbb{F}$ and $a\ge 0$ and $b\ge 0$ then $a+b\ge 0$.

(f) If $a,b\in\mathbb{F}$ and $0<a<b$ then $b^{-1}<a^{-1}$.

Proof.

(a) Assume $a\in\mathbb{F}$ and $a>0$.

Then $a+(-a)>0+(-a)$, by (OFb).

So $0>-a$, by (Fd) and (Fc).

(b) Assume $a\in\mathbb{F}$ and $a\ne 0$.

Case 1: $a>0$.

Then $a\cdot a>a\cdot 0$, by (OFb).

So $a^2>0$, by part (a).

Case 2: $a<0$.

Then $-a>0$, by part (a).

Then $(-a)^2>0$, by Case 1.

So $a^2>0$, by Proposition 26.3 (f).

(c) To show: $1 \geq 0$.
$1 = 1^2 \geq 0$, by part (b).

(d) Assume $a \in \mathbb{F}$ and $a > 0$.
By part (b), $a^{-2} = (a^{-1})^2 > 0$.
So $a(a^{-1})^2 > a \cdot 0$, by (OFb).
So $a^{-1} > 0$, by (Fh) and Proposition 26.3 (a).

(e) Assume $a, b \in \mathbb{F}$ and $a \geq 0$ and $b \geq 0$.

$$\begin{aligned} a + b &\geq 0 + b, \quad \text{by (OFa)}, \\ &\geq 0 + 0, \quad \text{by (OFa)}, \\ &= 0, \quad \text{by (Fc)}. \end{aligned}$$

(f) Assume $a, b \in \mathbb{F}$ and $0 < a < b$.
So $a > 0$ and $b > 0$.
Then, by part (d), $a^{-1} > 0$ and $b^{-1} > 0$.
Thus, by (OFb), $a^{-1}b^{-1} > 0$.
Since $a < b$, then $b - a > 0$, by (OFa).
So, by (OFb), $a^{-1}b^{-1}(b - a) > 0$.
So, by (Fh), $a^{-1} - b^{-1} > 0$.
So, by (OFa), $a^{-1} > y^{-1}$.

□

Proposition 26.5. *Let $\mathbb{F}$ be an ordered field and let $x, y \in \mathbb{F}$ with $x \geq 0$ and $y \geq 0$. Then*

$$x \leq y \qquad \textit{if and only if} \qquad x^2 \leq y^2.$$

Proof. Assume $x, y \in S$ and $x \geq 0$ and $y \geq 0$.
To show: (a) If $x \leq y$ then $x^2 \leq y^2$.
(b) If $x^2 \leq y^2$ then $x \leq y$.

(b) Assume $x^2 \leq y^2$.
Adding $(-x^2)$ to each side and using (OFa) gives $y^2 + (-x^2) \geq x^2 + (-x^2) = 0$.
So $y^2 - x^2 \geq 0$.
Using Proposition 26.3(e) and axioms (Ff) and (Fi),

$$\begin{aligned} (y - x)(y + x) &= yy + (-x)y + yx + (-x)x = y^2 + (-xy) + xy + (-xx) \\ &= y^2 + 0 - x^2 = y^2 - x^2. \end{aligned}$$

So $(y - x)(y + x) \geq 0$.
By Proposition 26.4(e) and Proposition 26.4 (d),
since $x \geq 0$ and $y \geq 0$ then $x + y \geq 0$ and $(x + y)^{-1} > 0$ (or $x = 0$ and $y = 0$).
So, by (OFb), $(y - x)(y + x)(x + y)^{-1} \geq 0$.
Using (Fg), then $y - x \geq 0$.
Adding x to both sides and using (OFa) gives $y \geq x$.

(a) Assume $y \geq x$.
Then $y - x \geq 0$.

Since $y \geq 0$ and $x \geq 0$ then, by (OFa), $(y+x) \geq y+0 = y \geq 0$.
So, by (OFb), $(y-x)(y+x) \geq 0$.
So $y^2 - x^2 \geq 0$.
So $y^2 \geq x^2$.

□